



Quelques aspects algorithmiques sur les systèmes de fermeture

Yoan Renaud

► To cite this version:

Yoan Renaud. Quelques aspects algorithmiques sur les systèmes de fermeture. Algorithmes et structure de données [cs.DS]. Université Blaise Pascal - Clermont-Ferrand II, 2008. Français. NNT : 2008CLF21897 . tel-00731341

HAL Id: tel-00731341

<https://theses.hal.science/tel-00731341>

Submitted on 12 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre : D.U. 1622

EDSPIC : 341

Université Blaise Pascal Clermont-Ferrand II

ECOLE DOCTORALE

SCIENCES POUR L'INGÉNIEUR DE CLERMONT-FERRAND

THÈSE

présentée par

Yoan RENAUD

pour obtenir le grade de

DOCTEUR D'UNIVERSITÉ

Spécialité : INFORMATIQUE

Quelques aspects algorithmiques sur les systèmes de fermeture

Soutenue publiquement le 8 décembre 2008 devant le jury :

Pr. Bernard MONJARDET	Rapporteur et Examineur
Dr. Vincent DUQUENNE	Rapporteur et Examineur
Pr. Alain QUILLIOT	Examineur
Pr. Michel HABIB	Examineur
Pr. Jean-Francois BOULICAUT	Examineur
Dr. Sylvie GUILLAUME	Examineur
Pr. Lhouari NOURINE	Directeur de Thèse

Mis en page avec la classe thloria.

Remerciements

La conclusion d'une thèse se fait généralement par les remerciements, c'est de coutume. Mais plus qu'une coutume, cela permet de mettre en évidence le fait qu'une thèse ne peut être menée à bon terme sans les différentes personnes qui nous entourent tout au long de ces années.

Dans un premier temps, je tiens à remercier Lhouari Nourine qui m'a permis de traiter des problèmes divers dans des domaines variés de l'informatique tous aussi intéressants les uns que les autres.

Je voudrais aussi remercier Vincent Duquenne et Bernard Monjardet pour avoir accepté d'être rapporteur de ma thèse ainsi que pour leurs divers conseils et remarques qui furent très enrichissants. Mes remerciements se dirigent aussi vers les différents membres qui ont composés mon jury : Michel Habib, Jean-François Boulicaut, Alain Quilliot et Sylvie Guillaume.

Durant ma thèse, j'ai aussi pu côtoyer diverses personnes. Les premières personnes que je voudrais remercier sont les différents membres de l'équipe avec laquelle j'ai eu le privilège de travailler. Raoul Medina et Olivier Raynaud m'ont ainsi accueilli chaleureusement et m'ont aidé à m'intégrer dans ce groupe de travail. Ils ont toujours répondu présent pour m'aider dans mes différents travaux de recherche et d'enseignement.

Différentes rencontres ont jalonné mes recherches qui ont permis un enrichissement tant sur le plan personnel que professionnel et je remercie toutes ces personnes. Ainsi, je tiens plus particulièrement à remercier Rokia Missaoui avec qui j'ai travaillé sur des problématiques de ma thèse. Les différentes séances de travail et discussions que nous avons eues ont permis d'atteindre une certaine maturité dans mon travail. Ce fut aussi très agréable de travailler avec elle de part sa disponibilité et sa gentillesse.

Bien sur, il est indispensable de remercier mes collègues et amis de bureau. Le bureau D009 fut pour moi ma deuxième maison durant ces années de thèse. Je tiens ainsi à remercier les différents locataires de ce bureau. Merci à Julien Ponge, Alain Gély (parti vers de nouveaux horizons à Metz) et Olivier Coupelon. Les divers moments que nous avons passés ont été très enrichissants. Ce fut très agréable de travailler (mais aussi se divertir) avec des personnes avec qui le courant est tout de suite passé. Bien sur je n'oublie pas Pierre Colomb et Ramy Ragab du bureau B118 qui nous rendent visite régulièrement à notre bureau et avec qui j'ai passé de superbes moments tant pour le travail qu'en dehors. La vie au laboratoire n'aurait pas été ce qu'elle a été sans Béatrice Bourdieu, Françoise Toledo et Rose Séguy qui ont toujours répondu présentes tout au long de ma thèse. Elles représentent le rayon de soleil de ce laboratoire.

Il ne faut pas oublier aussi qu’une thèse n’arrive qu’après cinq années de cursus universitaire. Ainsi, je tiens à remercier toutes les personnes que j’ai pu côtoyer durant ces diverses années. Mes remerciements vont au “club des cinq” en les personnes de Sylvain Vacheron, Cédric Masson, Thomas Busy et Cédric Léonard et moi même avec lesquels j’ai passé mes différentes années à la fac avec grand plaisir et beaucoup de moments uniques pendant et en dehors des cours (Au nota par exemple). Je voudrais remercier aussi Bibine, Stouf, Bachelor et tous les autres qui se reconnaîtront qui ont jalonné mon parcours universitaire.

Enfin, mes derniers remerciements, mais pas des moindres, vont à ma famille et à mes proches. Un GRAND MERCI à mon père et ma mère Luc et Jeannette qui m’ont toujours épaulés dans mes choix et soutenus dans les moments difficiles. Bien sur, ce grand merci va aussi à ma soeur Stéphanie et toute sa petite famille : Manu, Tess et Ilona pour tout ce qu’ils ont pus faire pour moi depuis des années. Je finirais par les personnes qui me sont chères et proches dans mon entourage. Un gros merci à Jacques qui m’a épaulé et soutenu tout au long de ce travail. Un grand merci à Michelle et Gilles, Serge, Martine et Lola, Sylvain et Nathalie ainsi que toutes les autres personnes qui me sont proches pour leur présence et leur soutien.

*“L’informatique n’est pas plus la science des ordinateurs
que l’astronomie n’est celle des télescopes.”*

*“Computer science is no more about computers
than astronomy is about telescopes.”*

Edsger Dijkstra.

Table des matières

I	Introduction	1
II	Systèmes de fermeture et théories de Horn : Définitions et relations	9
1	Systèmes de fermeture	11
1.1	Ordre partiel et treillis	11
1.2	Systèmes de fermeture	13
1.3	Treillis des systèmes de fermeture	15
1.4	Représentation des systèmes de fermeture	17
1.4.1	Contextes formels	17
1.4.2	Systèmes d'implications	19
1.5	Récapitulatif	23
2	Théories de Horn : Relation avec les systèmes de fermeture	27
2.1	Définitions	27
2.2	Formes Normales	29
2.3	Fonctions de Horn	30
2.3.1	FNC de Horn et bases d'implications	30
2.3.2	Modèles caractéristiques et éléments inf-irréductibles	31
<hr/>		
III	Différentes opérations sur les systèmes de fermeture	35
	Introduction	37
3	Borne inférieure de systèmes de fermeture	43
3.1	Éléments inf-irréductibles	43

3.2	Base d'implications	46
3.3	Opérateur de fermeture	48
3.4	Conclusion	49
4	Borne supérieure de systèmes de fermeture	51
4.1	Éléments inf-irréductibles	51
4.2	Base d'implications	53
4.2.1	Cas particulier : bases d'implications directes	54
4.2.2	Cas général	57
4.3	Opérateur de fermeture	58
4.4	Conclusion	59
5	Différence de deux systèmes de fermeture	61
5.1	Éléments inf-irréductibles	61
5.2	Base d'implications	64
5.3	Opérateur de fermeture	65
5.4	Conclusion	66

IV	Problème de génération des bases d'implications mixtes	69
	Introduction	71
6	Définitions et Problématique	75
6.1	Définitions et notations	75
6.1.1	Contexte complémentaire et apposition de contextes	75
6.1.2	Implications positives, négatives et mixtes	76
6.2	Problématique	78
7	Génération d'implications mixtes	83
7.1	Propriétés et règles d'inférence sur les implications mixtes	84
7.1.1	Propriétés	84
7.1.2	Règles d'inférence	86
7.1.3	Récapitulatif	89
7.2	Complétude de la génération des implications mixtes	90
7.3	Cas particulier : Contexte réduit	92
7.3.1	Contexte complet	93
7.3.2	Nouvelle règle d'inférence	95
7.3.3	Complétude de la génération d'implications mixtes	96
7.3.4	Exemple complet	100
7.4	Conclusion	102
V	Conclusion	103
8	Conclusion et Perspectives	105
	Bibliographie	109

Première partie

Introduction

Depuis des millénaires, l'homme a cherché à résoudre de nombreux problèmes. La conclusion de ces recherches était, la plupart du temps, infructueuse et les objectifs durs, voir impossible à atteindre sans méthode rigoureuse. L'être humain a ainsi créé et utilisé des outils l'aidant à calculer et à automatiser diverses tâches afin d'aboutir à une solution ou de montrer que celle-ci n'est pas atteignable.

Ainsi la logique a fait son apparition, et on a assisté à un développement fulgurant de l'approche mathématique et informatique de la logique au courant du vingtième siècle. L'électronique et l'apparition de l'ordinateur permirent ainsi d'automatiser les différents calculs à effectuer et d'accélérer considérablement le temps menant à la résolution d'un problème.

Mais malgré les avancées technologiques et l'augmentation permanente de puissance des machines, certains problèmes nécessitent toujours un temps de calcul qui n'est pas envisageable en terme de traitement pratique. Un des facteurs prépondérants est l'augmentation incessante de la taille des données qui doivent être traitées. C'est ce que l'on appelle l'explosion combinatoire. Il suffit de considérer un ensemble de données de cardinalité n pour constater que l'ensemble des parties comporte 2^n éléments ce qui représente un nombre exponentiel de configurations possibles et donc un ensemble très large de possibilités de réponse et de solutions à tester.

De ce fait, de nombreux problèmes de recherche restent encore à ce jour ouverts, c'est à dire ne possèdent pas d'algorithme efficace permettant d'arriver à la solution souhaitée.

Les chercheurs étudient ces problèmes à la recherche de méthodes et de structures mathématiques permettant d'aboutir à des algorithmes pertinents ou bien au contraire de démontrer que le problème étudié n'est pas traitable.

Depuis plus d'un demi siècle, les études sur les structures d'ordres [24] et la logique se sont multipliées afin de modéliser et traiter les problèmes dans divers domaines tel que l'intelligence artificielle, les bases de données ou encore le datamining pour ne citer qu'eux. Elles sont devenues prépondérantes dans la validation et la résolution de beaucoup de problèmes.

Les treillis et les systèmes de fermeture représentent des structures algébriques et combinatoires qui sont des outils de modélisation naturels [14, 17, 18, 25, 26, 28, 54, 75, 101]. Nous retrouvons fréquemment en informatique des relations sur les objets qui peuvent ainsi être modélisées par ces outils.

De ce fait, plusieurs représentations ont été étudiées pour appréhender efficacement les

systèmes de fermeture mais aussi pour qu'ils soient un outil efficace dans divers domaines. Parmi celles-ci, nous pouvons distinguer l'opérateur de fermeture associé à un système de fermeture \mathcal{F} , qui, appliqué à un ensemble A , renvoie le plus petit ensemble fermé de \mathcal{F} contenant A . L'ensemble des éléments inf-irréductibles d'un système de fermeture le caractérise de la même manière. Ainsi un système de fermeture peut être construit à partir de la fermeture par intersection de ses éléments inf-irréductibles. Une autre représentation, très importante de par ses applications, est le système d'implications complet ou base d'implications, lié à un système de fermeture. Cette base d'implications associée à un système de fermeture permet de reconstituer les ensembles fermés et met aussi en relation de cause à effet les éléments de l'ensemble sur lequel est construit le système de fermeture.

Nous nous sommes ainsi intéressés pendant cette thèse aux aspects algorithmiques et structurelles sur les systèmes de fermeture et aux répercussions sur leurs différentes représentations.

Un premier travail fut réalisé au début de cette thèse avec la collaboration de Alain Gély. Le but de ce travail était de caractériser la famille des systèmes de fermeture partageant le même ensemble d'implications non unitaires (implications du types $X \rightarrow Y$ tel que $|X| \geq 1$) dans leur base de Guigues Duquenne. Nous avons caractérisé le plus grand système de fermeture partageant les mêmes implications non unitaires dans leur base d'implications de Guigues-Duquenne [38]. Nous avons ainsi défini l'opération d'atomisation qui permet d'obtenir un tel système de fermeture. Ce travail avait fait état d'une publication à la conférence ICFCA05 [57]. Nous ne le développerons pas dans ce manuscrit puisqu'il a déjà été traité dans le manuscrit de thèse de Alain Gély. Divers travaux autour des aspects algorithmiques sur les systèmes de fermeture ont ainsi été effectués au sein de notre équipe de recherche [76, 82, 83].

Nous avons continué nos travaux sur différents aspects et représentations des systèmes de fermeture. Ce manuscrit de thèse se décompose en trois grandes parties qui sont détaillées ci-dessous.

II. Systèmes de fermeture et fonctions de Horn : Définitions et relations

La deuxième partie (chapitre 1 et 2) est une introduction générale sur les systèmes de fermeture et les fonctions de Horn en logique propositionnelle.

Le chapitre 1 donne les définitions sur les systèmes de fermeture et présente les différentes représentations qui peuvent lui être associées. Il se décompose en cinq sections.

Dans la première section, nous donnons les définitions des ordres partiels et des treillis.

Les systèmes de fermeture et les treillis sont des structures mathématiques fortement liées. Nous donnons ainsi les différentes définitions sur les systèmes de fermeture dans la deuxième section et définissons le treillis des systèmes de fermeture dans la section 3.

La quatrième section présente différentes représentations des systèmes de fermeture en s'attardant plus particulièrement sur les bases d'implications.

Enfin une dernière section rappelle les différents liens qui existent entre les différentes notions exposées précédemment.

Le chapitre 2 traite des théories de Horn en logique propositionnelle et de leurs liens avec les systèmes de fermeture.

Dans un premier temps, nous donnons les définitions en logique propositionnelle nécessaires afin d'appréhender les théories de Horn. Puis nous exposons la bijection qui existe entre les théories de Horn et les systèmes de fermeture.

III. Différentes opérations sur les systèmes de fermeture

La partie 3 (chapitre 3 à 6) présente trois opérations ensemblistes sur les systèmes de fermeture.

- La borne inférieure de systèmes de fermeture.
- La borne supérieure de systèmes de fermeture.
- La différence de deux systèmes de fermeture.

Outre un intérêt théorique, ces trois opérations sont appliqués et applicables dans différents domaines tel que l'intelligence artificielle, la logique, les bases de données et bien d'autres. Les représentations des systèmes de fermeture considérés peuvent ainsi être différentes selon les problèmes auxquels nous les rattachons. Nous étudions donc ces trois opérations sur trois représentations des systèmes de fermeture que sont :

- Les éléments inf-irréductible.
- La base d'implications.
- L'opérateur de fermeture.

Le but de cette partie est donc d'étudier ces trois opérations en considérant les différentes représentations des systèmes de fermeture évoquées ci-dessus.

L'introduction de cette partie expose ces trois opérations et les met en relation avec leurs problèmes équivalents étudiés sur les théories de Horn en logique propositionnelle [41, 42, 44].

Dans le chapitre 4, nous étudions l'opération de borne inférieure de systèmes de fermeture. Cette opération est équivalente à l'opération d'intersection sur les théories de Horn. Nous transcrivons ainsi les résultats exposés dans [42] sur les systèmes de fermeture. De plus, nous définissons un opérateur de fermeture associé à la borne inférieure de systèmes de fermeture qui utilise uniquement les opérateurs de fermeture de chaque système de fermeture considéré.

Le chapitre 5 traite de l'opération de borne supérieure de systèmes de fermeture. Nous avons ainsi remarqué que cette opération est équivalente au calcul de l'enveloppe de Horn de la disjonction de théories de Horn qui a été étudié dans [41]. Nous exposons ainsi leurs résultats en les exprimant sur les systèmes de fermeture. Nous nous attardons plus particulièrement sur la représentation des systèmes de fermeture par les bases d'implications. Il a été montré dans [41] qu'il n'existe pas d'algorithme polynomial permettant de calculer une base d'implications de la borne supérieure de systèmes de fermeture étant données leurs bases d'implications respectives. Une contribution de cette thèse est d'avoir montré que ce problème est polynomial si nous considérons que les bases d'implications des systèmes de fermeture considérés sont directes. L'algorithme qui en découle a fait l'objet d'une publication lors de la conférence MCO'08 [89].

Le chapitre 6 traite de l'opération de la différence de deux systèmes de fermeture. Ce problème étant équivalent au problème de différence de deux théories de Horn qui a été étudié dans [44], nous avons retranscrit leurs résultats sur les systèmes de fermeture. Cette dernière opération n'est pas défini sur les systèmes de fermeture mais peut avoir des applications dans divers domaines.

IV. Problème de génération des bases d'implications mixtes

Enfin, la quatrième partie de cette thèse (chapitre 6 et 7) porte sur le problème de génération des bases d'implications mixtes. Nous avons vu qu'une base d'implications est une représentation des systèmes de fermeture très utilisée dans différents domaines. Mais celle-ci permet seulement des associations positives entre les éléments. Le but

de cette partie est d'étudier les bases d'implications prenant en compte à la fois la présence et l'absence d'éléments dans les implications. Nous appelons ces implications des "implications mixtes".

Le chapitre 6 présente ainsi les notations et les définitions que nous allons utiliser pour les bases d'implications mixtes. Nous introduisons une méthode naïve de générations des implications à partir d'un contexte formel. Nous regardons ensuite différentes problématiques liées à la génération des bases d'implications purement positives, purement négatives et mixtes.

Dans le chapitre 7, nous nous concentrons plus particulièrement sur le problème de la génération d'une base d'implications mixtes à partir des bases d'implications purement positives et purement négatives associées à un contexte formel R . Nous verrons ainsi, qu'à partir de cette information, nous n'avons aucune garantie d'obtenir une base d'implications mixtes complète pour un contexte R considéré. Cependant, nous exhiberons un certain nombre de propriétés et de règles d'inférence qui nous permettent de déduire certaines implications mixtes de R . Ce travail a fait état d'une publication lors de la conférence ICFCA 08 [78].

Nous montrerons aussi que dans le cas particulier d'un contexte réduit, nous obtenons une nouvelle règle d'inférence qui nous permet de démontrer la justesse et la complétude de la base d'implications mixtes déduite de toutes les règles d'inférences que nous avons trouvées et ceci uniquement à partir des bases d'implications positives et négatives.

Enfin un chapitre de conclusion permettra de faire le bilan de cette thèse et proposera quelques perspectives de recherche.

Deuxième partie

Systèmes de fermeture et théories de Horn : Définitions et relations

Chapitre 1

Systèmes de fermeture

Sommaire

1.1	Ordre partiel et treillis	11
1.2	Systèmes de fermeture	13
1.3	Treillis des systèmes de fermeture	15
1.4	Représentation des systèmes de fermeture	17
1.4.1	Contextes formels	17
1.4.2	Systèmes d'implications	19
1.5	Récapitulatif	23

1.1 Ordre partiel et treillis

Dans cette section, nous rappelons quelques définitions sur les ordres partiels et les treillis. Les différentes notations et définitions que nous utilisons sont tirées des différents ouvrages connus sur ces notions [28, 54, 24].

Définition 1 *Ordre partiel*

Soit G un ensemble. Un ordre partiel sur G est une relation binaire \leq sur G telle que, pour tout $x, y, z \in G$,

- (i) $x \leq x$ (réflexivité)
- (ii) $x \leq y$ et $y \leq x$ implique $x = y$ (antisymétrie)
- (iii) $x \leq y$ et $y \leq z$ implique $x \leq z$ (transitivité)

Un ordre partiel sur un ensemble G est noté (G, \leq) .

Définition 2 *Relation de couverture*

Soit (G, \leq) un ordre partiel et $x, y \in G$. On dit que x est couvert par y (ou y couvre x) et on note $x \prec y$ si $x < y$ et $x \leq z < y$ implique $z = x$. Autrement dit, il n'existe pas $z \in G$ avec $x < z < y$.

Définition 3 *Idéal et filtre*

Soit (G, \leq) un ordre partiel et $A \subseteq G$,

- A est un idéal si, pour tout $x \in A$, $y \in G$ et $y \leq x$, on a $y \in A$.
- A est un filtre si, pour tout $x \in A$, $y \in G$ et $y \geq x$, on a $y \in A$.

Pour $j \in G$, on notera l'idéal (resp. le filtre) principal de j par $\downarrow j = \{x \in G \mid x \leq j\}$ (resp. $\uparrow j = \{x \in G \mid x \geq j\}$). Pour $A \subseteq G$, on notera $\downarrow A = \bigcup_{j \in A} \downarrow j$ (resp. $\uparrow A = \bigcup_{j \in A} \uparrow j$) l'idéal (resp. le filtre) associé à A .

Définition 4 *Borne supérieure, borne inférieure*

Considérons un ordre partiel (G, \leq) . Un élément $x \in G$ est la borne supérieure (ou supremum) de $y, z \in G$ si les deux propriétés suivantes sont vérifiées :

1. $y \leq x$ et $z \leq x$, et
2. Pour tout $t \in G$ tel que $y \leq t$ et $z \leq t$, on a $x \leq t$.

On définit de façon duale la borne inférieure (ou infimum).

On note par $x \vee y$ (resp. $x \wedge y$) le supremum (resp. l'infimum) de x et y .

Définition 5 *Demi-treillis et treillis*

Soit $T = (G, \leq)$ un ordre partiel non vide. T est un inf-demi-treillis si toute paire $\{x, y\}$ de ses éléments admet un infimum $x \wedge y$. C'est un sup-demi-treillis si toute paire $\{x, y\}$ de ses éléments admet un supremum $x \vee y$. T est un treillis s'il est à la fois inf-demi-treillis et sup-demi-treillis.

Un treillis T possède un plus grand élément, noté \top , ainsi qu'un plus petit élément, noté \perp .

La figure 1.1 présente différents ordres partiels. La figure a) représente un treillis puisque toute paire de ses éléments $\{x, y\}$ admet un infimum $x \wedge y$ et un suprémum $x \vee y$. L'ordre partiel b) n'est donc pas un treillis puisque qu'il n'admet pas ces propriétés. En revanche l'ordre partiel c) représente un inf-demi-treillis puisque tout couple d'éléments admet un infimum.

Définition 6 *Atomes et éléments irréductibles :*

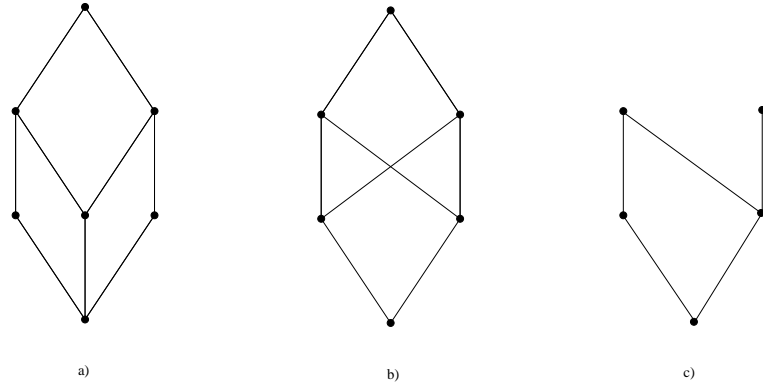


FIG. 1.1 – a) Treillis , b) Un ordre partiel, c) Un demi-treillis

- Les éléments couverts par \top se nomment les co-atomes du treillis.
- Les éléments qui couvrent \perp se nomment les atomes du treillis.
- $j \in T$ est dit \vee -irréductible si quels que soient $x, y \in T$, $x \vee y = j \Rightarrow x = j$ ou $y = j$.
- $m \in T$ est dit \wedge -irréductible si quels que soient $x, y \in T$, $x \wedge y = m \Rightarrow x = m$ ou $y = m$.

On peut aussi noter que les éléments \vee -irréductibles sont les éléments qui couvrent un seul élément et les éléments \wedge -irréductibles sont les éléments qui sont couverts par un seul élément.

1.2 Systèmes de fermeture

Nous allons voir à présent les différentes définitions sur les systèmes de fermeture. Nous montrons que les systèmes de fermetures, les opérateurs de fermeture et les treillis sont des structures mathématiques fortement liées. Un système de fermeture ordonné par inclusion est un treillis. De plus, à tout treillis correspond un système de fermeture et par conséquent un opérateur de fermeture.

Définition 7 Système de fermeture

Soit G un ensemble d'éléments. Une famille d'ensembles \mathcal{F} sur G est un système de fermeture si :

- a) $G \in \mathcal{F}$.
- b) $X, Y \in \mathcal{F}$ implique $X \cap Y \in \mathcal{F}$.

Si $F \in \mathcal{F}$, alors F est dit \mathcal{F} -fermé ou simplement fermé.

Remarque : Une famille qui vérifie la propriété b) de la définition 7 est appelée une théorie de Horn en logique. Si elle respecte les deux propriétés, elle est appelée fonction de Horn stricte (pour plus de détails, se référer au Chapitre 2).

Le terme *famille de Moore* est aussi usuellement utilisé pour un système de fermeture. Considérons l'exemple suivant d'un système de fermeture :

Exemple 1 Soit $G = \{a, b, c, d\}$ un ensemble d'éléments. Alors la famille d'ensembles $\mathcal{F} = \{\emptyset, a, b, c, ab, ad, bc, abcd\}$ est un système de fermeture sur G puisque tout ensemble peut être obtenu par intersection des ensembles le contenant.

Il existe une correspondance entre la notion de système de fermeture et celle d'opérateur de fermeture définie comme suit :

Définition 8 Opérateur de fermeture

Soit G un ensemble d'éléments. L'application $\varphi : 2^G \rightarrow 2^G$ est un opérateur de fermeture si les conditions suivantes sont vérifiées :

- $X \subseteq \varphi(X)$ (extension)
- $X \subset Y \Rightarrow \varphi(X) \subseteq \varphi(Y)$ (monotonie)
- $\varphi\varphi(X) = \varphi(X)$ (idempotence)

Un ensemble $X \subseteq G$ est dit fermé si $X = \varphi(X)$. L'ensemble des fermés par φ définit ainsi un système de fermeture [28].

On notera qu'un ensemble d'éléments inf-irréductibles est une représentation du système de fermeture associé. La fermeture de $\mathcal{M}(\mathcal{F})$ par intersection est ainsi égale à \mathcal{F} sachant que l'ensemble G appartient à \mathcal{F} . De manière plus formelle, nous pouvons le définir comme suit :

Proposition 1 Soit \mathcal{F} un système de fermeture sur G . Alors la famille des fermés définis par l'opérateur $\varphi_{\mathcal{F}}(A) = \bigcap \{M \in \mathcal{M}(\mathcal{F}) \mid A \subseteq M\}$ est égale à \mathcal{F} .

On notera par $\varphi_{\mathcal{F}}$ l'opérateur de fermeture associé au système de fermeture \mathcal{F} .

Proposition 2 Soit φ un opérateur de fermeture. $\mathcal{F}_{\varphi} = \{X \subseteq G \mid X = \varphi(X)\}$ est un système de fermeture.

Proposition 3 Soit \mathcal{F} un système de fermeture sur G . Alors (\mathcal{F}, \subseteq) est un treillis avec :

- $X \wedge Y = X \cap Y$
- $X \vee Y = \varphi_{\mathcal{F}}(X \cup Y) = \bigcap \{F \in \mathcal{F} \mid X \cup Y \subseteq F\}$

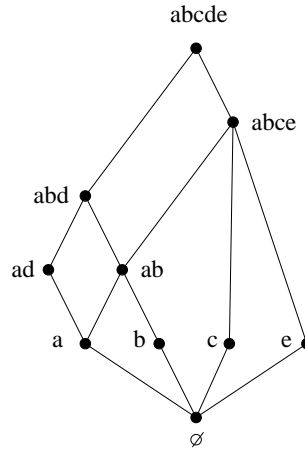


FIG. 1.2 – Treillis correspondant au système de fermeture de l'exemple 2.

Exemple 2 Soient $G = \{a, b, c, d, e\}$ et le système de fermeture sur G suivant :

$\mathcal{F} = \{\emptyset, \{a\}, \{b\}, \{c\}, \{e\}, \{ab\}, \{ad\}, \{abd\}, \{abce\}, \{abcde\}\}$.

Nous avons alors l'ensemble des éléments inf-irréductibles de \mathcal{F} : $\mathcal{M}(\mathcal{F}) = \{b, c, e, ad, abd, abce\}$ et l'ensemble des éléments sup-irréductibles de \mathcal{F} : $\mathcal{J}(\mathcal{F}) = \{a, b, c, e, ad\}$.

Le treillis correspondant à ce système de fermeture est présenté dans la figure 1.2.

1.3 Treillis des systèmes de fermeture

Nous avons vu précédemment la notion de système de fermeture et de treillis. Nous nous intéressons à présent à l'ensemble de tous les systèmes de fermeture définis sur un ensemble G que l'on note \mathcal{K} . Cet ensemble définit ainsi lui même un système de fermeture mais cette fois-ci sur l'ensemble 2^G .

Proposition 4 $\mathcal{K} = \{\mathcal{F} \subseteq 2^G \mid \mathcal{F} \text{ est un système de fermeture}\}$ est un système de fermeture avec $\varphi_{\mathcal{K}}(\mathcal{F}) = \bigcap \{\mathcal{F}' \in \mathcal{K} \mid \mathcal{F} \subseteq \mathcal{F}'\}$, où \mathcal{F} est une famille arbitraire sur G .

Corollaire 1 (\mathcal{K}, \subseteq) est un treillis avec pour tout $\mathcal{F}, \mathcal{F}' \in \mathcal{K}$ on a :

- $\mathcal{F} \wedge \mathcal{F}' = \mathcal{F} \cap \mathcal{F}'$.
- $\mathcal{F} \vee \mathcal{F}' = \varphi_{\mathcal{K}}(\mathcal{F} \cup \mathcal{F}') = \{F \cap F' : F, F' \in \mathcal{F} \cup \mathcal{F}'\}$.

La figure 1.3 illustre les opérations de l'infimum et du suprémum de deux systèmes de fermeture sur un ensemble G .

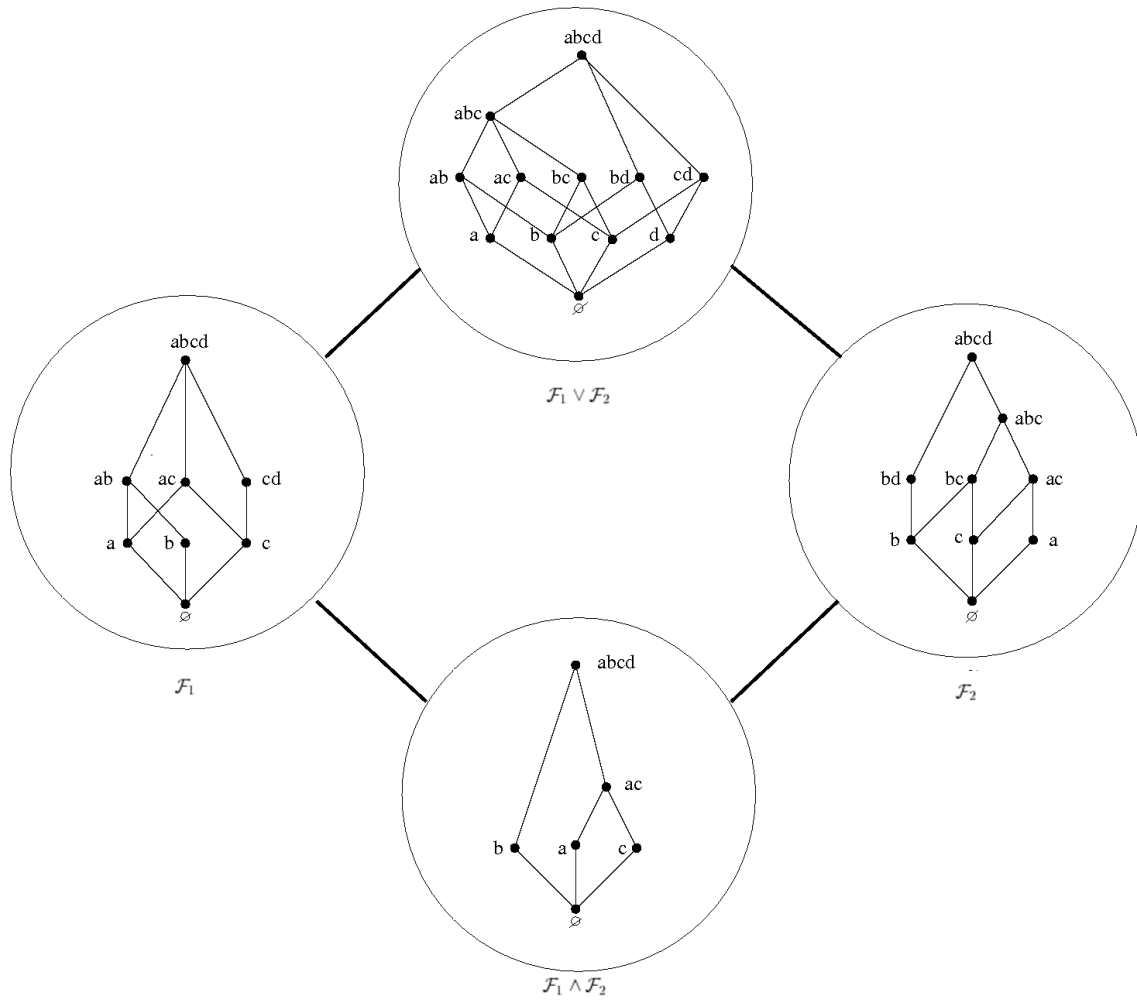


FIG. 1.3 – Opérations sur les systèmes de fermeture.

1.4 Représentation des systèmes de fermeture

Nous allons voir dans cette section les différentes représentations les plus utilisées pour les systèmes de fermeture. Nous avons déjà vu dans la section 1.2 la correspondance entre les systèmes de fermeture et les opérateurs de fermeture ainsi que les ensembles des éléments inf-irréductibles associés. Nous nous intéressons à présent à la représentation des systèmes de fermeture par les contextes formels et par les bases d'implications.

1.4.1 Contextes formels

La notion de contexte formel est une notion basique en analyse de concepts formels initiée par Ganter et Wille [54].

On note par $R = (G, M, I)$ un contexte formel où G représente l'ensemble des objets et M l'ensemble des attributs du contexte. I exprime la relation entre G et M . Dans le but d'exprimer le fait qu'un objet $g \in G$ est en relation I avec un attribut $m \in M$, on note gIm ou $(g, m) \in I$ qui est interprété comme "l'objet g contient l'attribut m ". La relation I est ainsi appelée relation d'incidence du contexte formel. On notera les objets du contexte par des nombres (1, 2, 3,...) et les attributs par des lettres (a, b, c,).

Définition 9 Soit $R=(G,M,I)$ un contexte formel.

On définit :

- $f : 2^G \rightarrow 2^M, f(X) = X' = \{a \in M \mid \forall x \in X, xIa\}$
- $g : 2^M \rightarrow 2^G, g(Y) = Y' = \{y \in G \mid \forall a \in Y, yIa\}$

Définition 10 Un concept formel d'un contexte $R = (G, M, I)$ est une paire (A, B) avec $A \subseteq G, B \subseteq M, A' = B$ et $B' = A$.

On note par $\mathcal{C}(R)$ l'ensemble de tous les concepts formels que l'on peut extraire d'un contexte formel R . Ces concepts sont ainsi partiellement ordonnés par :

$$(X_1, Y_1) \leq (X_2, Y_2) \Leftrightarrow X_1 \subseteq X_2, Y_1 \subseteq Y_2.$$

Un treillis des concepts associé à un contexte R est ainsi un treillis complet $(\mathcal{C}(R), \leq)$ que l'on notera par $\mathcal{B}(R)$.

Proposition 5 Soit $R = (G, M, I)$ un contexte. $\varphi_1 : 2^G \rightarrow 2^G$ avec $\varphi_1(A) = f \circ g(A)$ est un opérateur de fermeture sur G et $\varphi_2 : 2^M \rightarrow 2^M$ avec $\varphi_2(A) = g \circ f(A)$ est un opérateur de fermeture sur M . Les systèmes de fermeture associés sont ainsi les suivants :

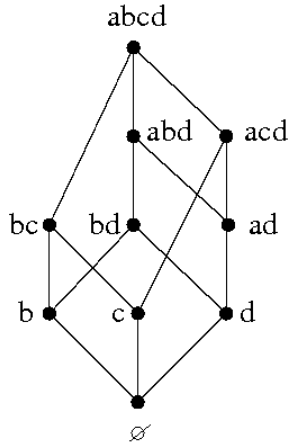
1. $\mathcal{F}_G = \{A \mid \varphi_1(A) = A\}$

$$2. \mathcal{F}_M = \{B \mid \varphi_2(B) = B\}$$

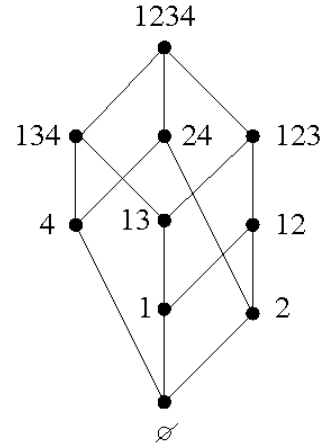
Nous pouvons voir sur la figure 1.4 un contexte formel représenté sous forme matricielle ainsi que les représentations des systèmes de fermeture associés sur les ensembles M et G du contexte R.

R	a	b	c	d
1	1	1	0	1
2	1	0	1	1
3	0	1	0	1
4	0	1	1	0

1) $R = (G, M, I)$



2) \mathcal{F}_M



3) \mathcal{F}_G

FIG. 1.4 – 1) Un contexte formel R, 2) Treillis représentant le système de fermeture \mathcal{F}_M , 3) Treillis représentant le système de fermeture \mathcal{F}_G .

Définition 11 *Un contexte $R = (G, M, I)$ est dit réduit si tout objet g de G ne peut être obtenu par intersection d'autres objets et si tout attribut m de M ne peut être obtenu par intersection d'autres attributs.*

Remarque : Nous considérons, dans la suite de ce manuscrit, qu'un contexte réduit ne contient pas d'objet composé de tous les attributs du contexte.

Proposition 6 [54] *Soient R un contexte formel et R' son contexte réduit associé. Alors les treillis associés à R et R' sont isomorphes.*

Il existe ainsi un unique contexte réduit associé à un contexte formel.

1.4.2 Systèmes d'implications

Une des représentations possibles pour un système de fermeture est une base d'implications. Nous allons voir dans cette section les différentes notations et définitions sur les bases d'implications ainsi que les relations existantes entre les notions de systèmes de fermeture et de bases d'implications.

Implications

Définition 12 Soit G un ensemble d'éléments. Un couple $(X, Y) \in 2^G \times 2^G$, noté $X \rightarrow Y$ est dit une implication sur G . X est appelé la prémisse de l'implication (ou encore l'antécédent), et Y la conclusion. On notera Σ une famille d'implications sur G .

Définition 13 Soient $X \rightarrow Y$ une implication et F un sous-ensemble de G . On dit que l'implication $X \rightarrow Y$ est satisfaite pour F (ou encore que F satisfait l'implication $X \rightarrow Y$) si et seulement si :

$$X \subseteq F \Rightarrow Y \subseteq F$$

Plus généralement, on dira qu'une famille d'implications est satisfaite pour un ensemble $F \subseteq G$ si et seulement si F satisfait toutes les implications de Σ .

Dans le cas où un ensemble F satisfait la famille d'implications Σ , on dira que F est un ensemble Σ -fermé. Dans le cas où F ne satisfait pas la famille d'implications Σ , le plus petit ensemble Σ -fermé contenant F est appelé la fermeture de F par Σ , notée F^Σ .

Étant donné un ensemble F , il existe des algorithmes permettant de calculer sa fermeture en temps linéaire par rapport à la taille de Σ [9, 71]. La taille d'une base d'implications est définie par $|\Sigma|$ qui correspond au nombre d'implications dans Σ .

Soit Σ un ensemble d'implications sur G . La famille des ensembles Σ -fermés constitue un système de fermeture, noté $\mathcal{F}_\Sigma = \{F^\Sigma \mid F \subseteq G\}$.

Dans le cas où plusieurs systèmes de fermetures \mathcal{F} , \mathcal{F}' , \mathcal{F}'' sont manipulés en même temps, on notera $\Sigma_{\mathcal{F}}$, $\Sigma_{\mathcal{F}'}$, $\Sigma_{\mathcal{F}''}$ les familles d'implications vérifiées correspondantes.

Propriété 1 Soient Σ une famille d'implications sur G et $X \rightarrow Y$ une implication. On dira que $X \rightarrow Y$ dérive de Σ (ou, de façon équivalente, que Σ infère $X \rightarrow Y$, ou encore que Σ implique logiquement $X \rightarrow Y$), si et seulement si $Y \subseteq X^\Sigma$.

On notera $\Sigma \vdash X \rightarrow Y$ l'implication logique de $X \rightarrow Y$ par Σ .

De la même façon qu'une implication dérive d'une famille d'implications Σ , on

dira qu'une famille d'implications Σ' dérive de Σ (noté $\Sigma \vdash \Sigma'$) si et seulement si, pour tout $X \rightarrow Y \in \Sigma'$, $\Sigma \vdash X \rightarrow Y$.

Définition 14 Deux familles d'implications Σ et Σ' sont équivalentes si et seulement si :

$$\Sigma' \vdash \Sigma \text{ et } \Sigma \vdash \Sigma'$$

ou

$$\mathcal{F}_\Sigma = \mathcal{F}_{\Sigma'}$$

A un système d'implications, on peut associer de façon unique un système de fermeture. En revanche, un système de fermeture peut correspondre à plusieurs familles d'implications. Dans ce cas, ces familles d'implications sont toutes *équivalentes*.

Base d'implications

Définition 15 Soit \mathcal{F} un système de fermeture. On dira que Σ est une base d'implications pour \mathcal{F} ssi $\mathcal{F} = \mathcal{F}_\Sigma$. On note par φ_Σ l'opérateur de fermeture associé à \mathcal{F}_Σ .

Les bases d'implications pour un système de fermeture peuvent être de taille exponentielle par rapport à la représentation que nous avons du système de fermeture auquel nous pouvons l'associer (par exemple les éléments inf-irréductibles de ce système de fermeture). Il est alors plus intéressant de se focaliser sur les plus petites d'entre elles en terme de nombre d'implications ou celles qui ont de bonnes propriétés.

Définition 16 Soient \mathcal{F} un système de fermeture et Σ une base d'implications de \mathcal{F} . On dira que Σ est

1. **Non redondante**, si pour tout $X \rightarrow Y \in \Sigma$, $\Sigma \setminus \{X \rightarrow Y\} \not\vdash X \rightarrow Y$.
2. **Minimum**, si pour tout Σ' une base équivalente à Σ , on a $|\Sigma'| \geq |\Sigma|$, où $|\Sigma|$ est le cardinal de Σ

De la définition ci-dessus, on déduit qu'une base minimum est une base non redondante, mais une base non redondante n'est pas forcément minimum.

Propriété 2 Axiomes d'Armstrong

Une base d'implications complètes associée à un système de fermeture sur G est un système d'implications qui vérifie les axiomes suivants :

1. $\forall A, B, C \subseteq G$, $A \rightarrow B$ et $B \rightarrow C$ impliquent $A \rightarrow C$ (transitivité)

2. $\forall A, B, Z \subseteq G, A \rightarrow B$ implique $AZ \rightarrow BZ$ (augmentation)
3. $\forall A, B \subseteq G, A \subseteq B$ implique $B \rightarrow A$. (réflexivité)

Les axiomes d'Armstrong [5] permettent ainsi de dériver toutes les implications qui sont satisfaites par un système de fermeture \mathcal{F} à partir d'une base d'implications de ce système de fermeture.

Parmi les différentes bases d'implications connues, nous nous intéressons plus particulièrement à trois bases d'implications que sont la base des générateurs minimaux ou base générique [85], la base des prémisses propres [101] et la base canonique plus communément appelée la base de Guigues Duquenne [38] ou bien encore la "stem base" dans [54]. Les deux premières bases d'implications évoquées sont des bases d'implications directes. On dira qu'une base d'implications est directe si pour tout ensemble $A \subseteq G$, il suffit d'un seul passage de Σ pour obtenir $\varphi_\Sigma(A)$. La base des générateurs minimaux est ainsi une base d'implications directe et la base des prémisses propres a la particularité d'être une base directe minimale.

Base des générateurs minimaux (générique)

La base des générateurs minimaux est construite à partir d'ensembles particuliers que sont les générateurs minimaux. Elle est uniquement définie mais n'a pas de propriété sur sa taille. De plus, cette base d'implications peut être redondante.

Définition 17 Soit \mathcal{F} un système de fermeture sur G . On appelle B un générateur minimal d'un ensemble fermé $F \in \mathcal{F}$ si $\nexists B' \subset F$ tel que $\varphi_{\mathcal{F}}(B') = F$.

Théorème 1 Soit \mathcal{F} un système de fermeture sur G . Alors $\Sigma_{\mathcal{F}} = \{B \rightarrow F \setminus B \mid B \subset G \mid B \text{ est un générateur minimal}\}$ est une base d'implications pour \mathcal{F} .

Exemple 3 Si nous considérons le système de fermeture de l'exemple 2, la base des générateurs minimaux associée à ce système de fermeture est la suivante :

$$\Sigma_{\mathcal{F}} = \{d \rightarrow a, bd \rightarrow a, ac \rightarrow be, bc \rightarrow ae, ae \rightarrow bc, be \rightarrow ac, ce \rightarrow ab, cd \rightarrow abc, de \rightarrow abc\}$$

Nous pouvons voir dans l'exemple 3, que cette base d'implications est redondante. Si nous considérons les règles $d \rightarrow a$ et $bd \rightarrow a$, nous remarquons que $bd \rightarrow a$ est redondante par rapport à $d \rightarrow a$. En général, cette base n'est donc pas forcément minimum.

Bases des prémisses propres [101]

Cette base d'implications est basée sur des éléments particuliers appelés "prémisses propres". Nous verrons que cette base est définie de manière unique et peut être redondante.

Définition 18 Soit un ensemble $A \subseteq G$, on note par :

$$A^\bullet = \varphi(A) \setminus (A \cup \bigcup_{x \in A} \varphi(A \setminus \{x\}))$$

l'ensemble des éléments de $\varphi(A)$ qui n'appartiennent ni à A ni à la fermeture d'aucun sous ensemble propre de A . On appelle A une **prémisse propre** si $A^\bullet \neq \emptyset$, c'est à dire :

$$\varphi(A) \neq A \cup \bigcup_{x \in A} \varphi(A \setminus \{x\})$$

En particulier, \emptyset est une prémisse propre si $\varphi(\emptyset) \neq \emptyset$.

Théorème 2 Soit \mathcal{F} un système de fermeture sur G . Alors $\Sigma_{\mathcal{F}} = \{A \rightarrow A^\bullet \mid A \text{ est une prémisse propre}\}$ est une base d'implications pour \mathcal{F} .

Exemple 4 Si nous considérons le système de fermeture de l'exemple 2, la base des prémisses propres associée à ce système de fermeture est la suivante :

$$\Sigma^{\mathcal{F}} = \{d \rightarrow a, ac \rightarrow be, bc \rightarrow ae, ae \rightarrow bc, be \rightarrow ac, ce \rightarrow ab, cd \rightarrow be, de \rightarrow bc\}$$

Nous pouvons voir dans cet exemple que l'ensemble des prémisses propres de cette base d'implications est un sous-ensemble de l'ensemble des générateurs minimaux associé au même système de fermeture.

Base de Guigues Duquenne [38]

La base de Guigues Duquenne [38] est construite sur des éléments particuliers appelés ensembles pseudo-fermés qui sont basés sur la définition des ensembles quasi-fermés. Elle est uniquement définie et a la propriété d'être minimum. En revanche, il n'est pas toujours possible de calculer la fermeture d'un ensemble en un seul passage de cette base d'implications.

Définition 19 Un ensemble $Q \subseteq G$ est un ensemble quasi-fermé si et seulement si $\mathcal{F} \cup Q$ est un système de fermeture.

Une caractérisation classique des ensembles quasi-fermés est la suivante :

Propriété 3 [38] Soit un ensemble $Q \subset G$. Q est quasi-fermé ssi pour tout $A \subset Q$, $A^{\Sigma^{\mathcal{F}}} = Q^{\Sigma^{\mathcal{F}}}$ ou $A^{\Sigma^{\mathcal{F}}} \subset Q$.

Enfin, à partir de la définition d'ensemble *quasi-fermé*, on peut donner la définition d'ensemble *pseudo-fermé*.

Définition 20 Un ensemble quasi-fermé P est un ensemble pseudo-fermé s'il n'existe pas d'ensemble quasi-fermé Q avec $Q \subset P$ et $Q^{\Sigma^{\mathcal{F}}} = P^{\Sigma^{\mathcal{F}}}$.

Exemple 5 Si nous considérons le système de fermeture de l'exemple 2, la base de Guigues Duquenne associée à ce système de fermeture est la suivante :

$$\Sigma^{\mathcal{F}} = \{d \rightarrow a, ae \rightarrow bc, be \rightarrow ac, ce \rightarrow ab, ac \rightarrow be, bc \rightarrow ae\}.$$

L'ensemble des pseudo-fermés permet ainsi de définir la base d'implications de Guigues Duquenne.

Théorème 3 Base minimum de Guigues-Duquenne[38].

Soit \mathcal{F} un système de fermeture, l'ensemble $\Sigma^{\mathcal{F}} = \{P \rightarrow P^{\Sigma^{\mathcal{F}}} \mid P \text{ est un ensemble pseudo-fermé}\}$ est une base d'implications minimum pour \mathcal{F} .

De plus, à partir d'une base d'implications quelconque, il existe un algorithme permettant de trouver la base de Guigues Duquenne associée. Cet algorithme a été découvert par Mayer puis Shock l'a simplifié [92].

Algorithme 1 : MINIMUM(Σ)

Données : Σ une base d'implications de \mathcal{F}

Résultat : Σ_{min} une base minimum d'implications de \mathcal{F}

début

pour chaque $X \rightarrow Y \in \Sigma$ **faire**

 Remplacer $X \rightarrow Y$ par $X \rightarrow X^{\Sigma}$;

pour chaque $X \rightarrow X^{\Sigma} \in \Sigma$ **faire**

si $X^{\Sigma^{\mathcal{F}}} \setminus \{X \rightarrow X^{\Sigma^{\mathcal{F}}}\} \in \mathcal{F}$ **alors** $\Sigma \setminus \{X \rightarrow Y\} \rightarrow \Sigma$

fin

1.5 Récapitulatif

Dans cette section, nous rappelons de manière brève les différents liens entre les notions définies dans ce chapitre.

Soit \mathcal{F} un système de fermeture sur G . Alors

- l'opérateur de fermeture $\varphi : 2^G \rightarrow 2^G$ avec $\varphi(X) = \bigcap \{F \in \mathcal{F} \mid X \subseteq F\}$ ou encore $\varphi(X) = \bigcap \{F \in \mathcal{F} \mid F \in \mathcal{M}(F), X \subseteq F\}$ correspond au système de fermeture \mathcal{F} .
- \mathcal{F} est isomorphe à la fermeture par intersection de la collection $\mathcal{M}(\mathcal{F}) \cup G$. Les éléments de $\mathcal{M}(\mathcal{F})$ peuvent donc servir à représenter le treillis.
- (\mathcal{F}, \subseteq) est un treillis. Notons que pour un treillis, on peut associer plusieurs systèmes de fermeture.
- Notons qu'on peut associer à \mathcal{F} plusieurs bases d'implications équivalentes, mais une seule est sous la forme de Guigues-Duquenne et une seule est sous la forme de base de générateurs minimaux.
- \mathcal{F} est égal à l'ensemble des ensembles Σ -fermés où Σ est une base de \mathcal{F} . Notons qu'une base minimum d'implications pour \mathcal{F} peut être calculée en temps polynomial à partir de \mathcal{F} . Par contre à partir de $\mathcal{M}(\mathcal{F})$, ce problème reste ouvert.

Les différentes représentations que nous avons vu dans ce chapitre sont donc très liés au système de fermeture correspondant. Malheureusement le passage entre ces différentes représentations reste, en général, assez difficile à effectuer. Le calcul d'un opérateur de fermeture pour un système de fermeture est une tâche facile si nous considérons son ensemble d'éléments inf-irréductibles ou bien une base d'implications associée. En revanche, concernant les autres représentations que nous avons étudiées, il n'existe pas à ce jour de méthode efficace permettant de passer de l'une à l'autre. Nous exhibons ces différents liens sur la figure 1.5 et regardons les proportions sur les tailles des représentations d'un système de fermeture dans le pire des cas sur la table 1.1.

	$\mathcal{M}(\mathcal{F})$	$\Sigma_{\mathcal{F}}$	\mathcal{F}
$ \mathcal{M}(\mathcal{F}) = m$	m	2^m	2^m
$ \Sigma_{\mathcal{F}} = m$	2^m	m	2^m
$ \mathcal{F} = m$	m	$ G \times m$	m

TAB. 1.1 – Taille des différentes représentations d'un système de fermeture \mathcal{F} sur G au pire des cas.

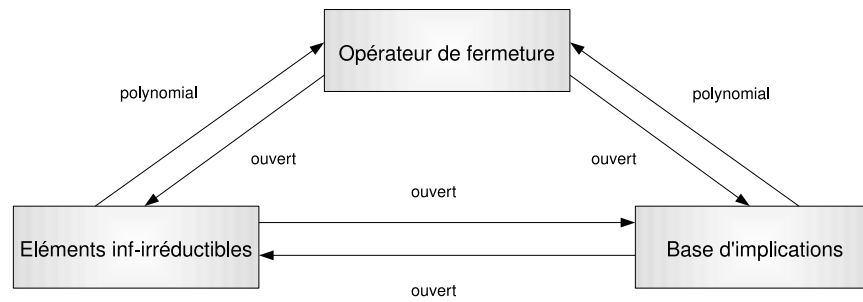


FIG. 1.5 – Passage entre les différentes représentations des systèmes de fermeture

Chapitre 2

Théories de Horn : Relation avec les systèmes de fermeture

Sommaire

2.1	Définitions	27
2.2	Formes Normales	29
2.3	Fonctions de Horn	30
2.3.1	FNC de Horn et bases d'implications	30
2.3.2	Modèles caractéristiques et éléments inf-irréductibles	31

Il existe une forte relation entre les systèmes de fermeture et les fonctions de Horn en logique propositionnelle. Dans ce chapitre, nous donnons les définitions et les notations nécessaires pour définir les fonctions de Horn. Ensuite, nous nous intéressons plus particulièrement aux relations qui existe entre les fonctions de Horn et les systèmes de fermeture.

2.1 Définitions

On note par $\{0, 1\}^n$ le produit cartésien $\prod_{i=1}^n \{0, 1\}$. Un tuple $v \in \{0, 1\}^n$ sera noté $v = (x_1, \dots, x_n)$ et v_i représente la i -ème composante, c'est à dire x_i . Si nous considérons un ensemble G tel que $|G| = n$, il y a une bijection entre $\{0, 1\}^n$ et $\mathcal{P}(G)$, l'ensemble des parties de G .

Définition 21 *Fonction booléenne*

On appelle fonction booléenne à n variables (ou fonction logique) toute application de la forme :

$$\begin{aligned} f : \{0,1\}^n &\rightarrow \{0,1\} \\ (x_1, \dots, x_n) &\rightarrow f(x_1, \dots, x_n) \end{aligned}$$

On peut décrire une fonction booléenne par une table de vérité ou bien utiliser des formules mathématiques à base d'opérateurs logiques. Ainsi le \wedge représente le "ET", \vee représente le "OU" et " \neg " représente le "NON".

L'étude des fonctions booléennes sur n variables est équivalente à l'étude de l'ensemble des parties de $\{0,1\}^n$ que l'on notera \mathcal{B}_n . On a de ce fait que $|\mathcal{B}_n| = 2^{2^n}$ où $|\mathcal{B}_n|$ représente le nombre de fonctions possibles sur n variables.

Nous pouvons ainsi représenter une famille \mathcal{F} d'ensembles de G par sa fonction caractéristique $f_{\mathcal{F}}$ comme suit :

$$f_{\mathcal{F}}(M) = \begin{cases} 1 & \text{si } M \in \mathcal{F} \\ 0 & \text{sinon} \end{cases}$$

De la même manière, on peut associer à une fonction booléenne f dans $\{0,1\}$, une famille d'ensembles sur G que l'on nomme les modèles de f .

$$\mathcal{F}_f = \{M \subseteq G \mid f(M) = 1\}$$

La bijection entre une fonction booléenne et une famille d'ensembles s'exprime de la manière suivante :

$$\mathcal{B}_n \rightarrow 2^{2^G}$$

Ainsi ces deux notations sont échangeables.

Définition 22 *Littéral :*

Un littéral est une fonction booléenne de la forme p (littéral positif) ou \bar{p} (littéral négatif), où p est une variable.

Définition 23 *Modèle et théorie :*

Une théorie \mathcal{T} est un sous-ensemble de $\{0,1\}^n$. Les éléments de \mathcal{T} sont appelés modèles.

Les modèles $(0,0,\dots,0)$ et $(1,1,\dots,1)$ seront respectivement notés par 0 et 1 .

Deux modèles v, w sont dits comparables (noté $v \leq w$) si pour tout $i \in \{1, \dots, n\}$, $v_i \leq w_i$ avec $0 \leq 1$.

Définition 24 *$T(f)$:*

Soit f une fonction booléenne, $T(f) = \{v \in \{0,1\}^n \mid f(v) = 1\}$ représente l'ensemble de tous les modèles de la fonction f .

On dira que deux fonctions booléennes f_1 et f_2 sont équivalentes si et seulement si $T(f_1) = T(f_2)$.

Exemple 6 La fonction $f = (a \wedge \bar{b}) \vee c$ est une fonction booléenne à trois variables.

$T(f) = \{(0, 0, 1), (1, 0, 0), (1, 0, 1), (0, 1, 1), (1, 1, 1)\}$.

La fonction $f' = (a \vee c) \wedge (\bar{b} \vee c)$ est équivalente à la fonction f puisque $T(f) = T(f')$

2.2 Formes Normales

En logique, les formes normales sont une normalisation des fonctions logiques. Il existe ainsi plusieurs types de formes normales. Nous nous intéressons ici plus particulièrement à la forme normale conjonctive et à la forme normale disjonctive.

Définition 25 *Clause et FNC :*

- Une clause est une fonction de la forme $c = l_1 \vee \dots \vee l_n$, $n \geq 0$, où chaque l_i est un littéral.
- Une fonction booléenne est dite en forme normale conjonctive (FNC) si et seulement si elle est de la forme $c_1 \wedge \dots \wedge c_m$, $m \geq 0$, où chaque c_i est une clause.

Définition 26 *Terme et FND :*

- Un terme est une fonction de la forme $t = l_1 \wedge \dots \wedge l_n$, $n \geq 0$, où chaque l_i est un littéral.
- Une fonction booléenne est dite en forme normale disjonctive (FND) si et seulement si elle est de la forme $t_1 \vee \dots \vee t_m$, $m \geq 0$, où chaque t_i est un terme.

Nous noterons la clause vide ($n = 0$) par \perp et le terme vide par \top . L'ensemble des littéraux positifs d'une clause c sera noté $P(c)$ et l'ensemble des littéraux négatifs par $N(c)$.

Théorème 4 *Soit f une fonction booléenne.*

- Il existe une fonction f_1 en FND telle que f_1 est équivalente à f .
- Il existe une fonction f_2 en FNC telle que f_2 est équivalente à f .

Le théorème 4 exprime le fait que l'on peut transformer une fonction f quelconque en une fonction ne contenant qu'une conjonction de clauses (FNC) ou bien une fonction ne contenant qu'une disjonction de termes.

Définition 27 *Un terme t est un impliquant d'une fonction booléenne f si t implique f , c'est à dire $\forall v \in \{0, 1\}^n$, $t(v) \leq f(v)$. Il est impliquant premier si tout sous-ensemble de conjonction de littéraux de t n'est pas un impliquant de f .*

Définition 28 Une clause c est un impliqué d'une fonction booléenne f si f implique c , c'est à dire $\forall v \in \{0,1\}^n, f(v) \leq c(v)$. Il est impliqué premier si tout sous-ensemble de disjonction de littéraux de c n'est pas un impliqué premier de f .

Exemple 7 Si nous considérons la fonction booléenne suivante :

$$f = (x_1 \wedge x_3) \vee (x_2 \wedge \overline{x_4}) \vee (x_3)$$

nous avons :

- $(x_1 \wedge x_3)$, $(x_2 \wedge \overline{x_4})$ et x_3 sont des impliquants
- $(x_1 \wedge x_3)$ n'est pas un impliquant premier.

De la même manière, si nous considérons la fonction booléenne qui suit :

$$f' = (x_1 \vee x_3) \wedge (x_2 \vee \overline{x_4}) \wedge (x_3).$$

- $(x_1 \vee x_3)$, $(x_2 \vee \overline{x_4})$ et (x_3) sont des impliqués.
- $(x_1 \vee x_3)$ n'est pas un impliqué premier.

2.3 Fonctions de Horn

Il existe des classes particulières pour les fonctions booléennes. Les fonctions booléennes de Horn forment ainsi une classe ayant des propriétés intéressantes que nous allons voir dans cette section. Il existe une forte relation entre les fonctions de Horn et les systèmes de fermeture. Plus exactement, les systèmes de fermeture sur un ensemble G et les fonctions de Horn strictes (c'est à dire qui peuvent être représentée par une conjonction de clauses strictes) sont en bijection.

2.3.1 FNC de Horn et bases d'implications

Définition 29 Une clause de Horn est une clause comportant au plus un littéral positif.

La dénomination "clause de Horn" vient du nom du logicien Alfred Horn qui mit en évidence l'intérêt de telles clauses en 1951 [60].

Il existe ainsi deux types de clauses de Horn distincts qui nous intéressent :

- **Les clauses de Horn strictes (ou pures)** : elles comportent un littéral positif et au moins un littéral négatif.
- **Les clauses de Horn négatives** : elles ne contiennent que des littéraux négatifs.

L'exemple 8 illustre ainsi les différents types de clauses de Horn.

Exemple 8

- $(a \vee \bar{b} \vee \bar{c})$ est une clause de Horn stricte.
- $(\bar{a} \vee \bar{b} \vee \bar{c})$ est une clause de Horn négative.
- $(a \vee b \vee \bar{c})$ n'est pas une clause de Horn.

Les clauses de Horn forment ainsi un sous-ensemble des formes normales disjonctives dans lesquels un seul terme est positif.

Définition 30 Une fonction booléenne est dite de Horn si elle est composée d'une conjonction de clauses de Horn.

L'ensemble des modèles d'une fonction de Horn est appelé une théorie de Horn. Si nous considérons que $v \wedge w$ correspond à l'opération “et” entre les deux vecteurs $v, w \in \{0, 1\}^n$ et que $Cl_\wedge(S)$ est la fermeture de $S \subseteq \{0, 1\}^n$ par \wedge , nous avons la propriété suivante :

Proposition 7 \mathcal{T} est une théorie de Horn si et seulement si $(\mathcal{T}) = Cl_\wedge(\mathcal{T})$.

Nous pouvons ainsi déduire de la proposition 7 que la théorie de Horn \mathcal{T} admet un unique modèle minimal $v = \bigwedge_{w \in \mathcal{T}} w$.

Théorème 5 Une fonction booléenne h de n variables $\{x_1, \dots, x_n\}$ est une fonction de Horn si et seulement si la famille composée de ses modèles et de $(1 \dots 1)$ forme un système de fermeture sur $G = \{x_1, \dots, x_n\}$.

Ce dernier théorème a été montré dans [60], évoqué par Dechter et Pearl dans [30] et utilisé dans [13] .

Il met ainsi en évidence l'équivalence entre les ensembles fermés d'un système de fermeture et les modèles d'une théorie de Horn. Nous avons vu que les clauses de Horn strictes sont de la forme $\bar{x}_1 \vee \dots \vee \bar{x}_k \vee y$. Ces clauses sont alors équivalentes à $\{x_1, \dots, x_k\} \vdash y$. Elles représentent donc des règles de type “si.... alors....”. Si nous les représentons sous forme implicationnelle, elles sont alors équivalentes à $x_1 \dots x_k \rightarrow y$ pour une base d'implications. Une FNC de Horn étant une conjonction de clauses de Horn, une représentation équivalente sera un ensemble d'implications.

2.3.2 Modèles caractéristiques et éléments inf-irréductibles

Définition 31 Un modèle v d'une théorie de Horn \mathcal{T} est appelé modèle caractéristique si $v \notin Cl_\wedge(\mathcal{T} \setminus \{v\})$.

Nous appellerons l'ensemble des modèles caractéristiques de \mathcal{T} l'ensemble caractéristique de \mathcal{T} que nous noterons $\mathcal{C}^*(\mathcal{T})$. Nous pouvons noter que chaque théorie de Horn admet un seul et unique ensemble de modèles caractéristiques et que les éléments maximaux d'une théorie appartiennent à son ensemble de modèles caractéristiques.

Remarque : Il existe différentes appellations de ces modèles dans la littérature. Nous avons vu la notion de “caractéristique” qui a été introduite dans [58] par Kautz, Kearns et Selman en 1993, mais il existe aussi le terme “extrême” introduit dans [31] par Dechter et Pearl en 1992. Dans la suite de ce manuscrit, nous utiliserons seulement la première notation afin d'éviter toute confusion.

Exemple 9 *Considérons les théories suivantes :*

$\mathcal{T}_1 = \{(000), (010), (001), (110), (101)\}$ et $\mathcal{T}_2 = \{(000), (010), (001), (110), (101), (100)\}$.

Si nous considérons $v = (110)$ et $w = (101)$, nous avons v et w qui appartiennent à \mathcal{T}_1 tandis que $v \wedge w = (100)$ n'appartient pas à \mathcal{T}_1 . Cela signifie donc que \mathcal{T}_1 n'est pas de Horn.

En revanche, nous avons $Cl_\wedge(\mathcal{T}_2) = \mathcal{T}_2$, ce qui signifie que \mathcal{T}_2 est une théorie de Horn.

Les modèles caractéristiques de \mathcal{T}_2 sont les suivants :

$$\mathcal{C}^*(\mathcal{T}_2) = \{(010), (001), (110), (101)\}$$

Si nous considérons ainsi les ensembles fermés du système de fermeture \mathcal{F} correspondants aux modèles caractéristiques de \mathcal{T} , cela signifie que ces ensembles ne peuvent pas être obtenus par intersection d'autres fermés de \mathcal{F} . Ces ensembles fermés correspondent donc aux éléments inf-irréductibles du système de fermeture \mathcal{F} plus G.

La table 2.1 présente les différentes équivalences entre les notions sur les systèmes de fermeture et celles sur les théories de Horn.

Théorie de Horn stricte	Système de fermeture
FNC de Horn	Base d'implications
Modèles caractéristiques	Eléments inf-irréductibles + G
Modèles	Ensembles fermés

TAB. 2.1 – Equivalences entre notions sur les systèmes de fermeture et les théories de Horn.

Troisième partie

Différentes opérations sur les systèmes de fermeture

Introduction

Les langages logiques sont souvent utilisés pour représenter la connaissance dans les bases de connaissances. La connaissance peut être ainsi consultée à l'aide de requêtes ce qui se traduira par une compilation de connaissance.

La compilation de connaissance (traduction littérale de “Knowledge compilation”) a pour but de mettre en évidence les inférences d’une manière plus efficace. La majorité des travaux qui ont été menés dans ce domaine se sont restreints à une représentation propositionnelle afin d’exprimer la connaissance de manière concise. Plus récemment, des travaux ont considéré une représentation basée sur des modèles afin, dans ce cas là, d’exprimer la connaissance de manière plus efficace. Si nous considérons le problème d’inférence propositionnelle, il consiste à vérifier si oui ou non $\Sigma \vdash \alpha$ où Σ et α représentent des fonctions propositionnelles et \vdash représente l’inférence logique. Malheureusement, dans le cadre de fonctions propositionnelles quelconques, ce problème est co-NP-complet et par conséquent non résoluble en temps polynomial hormis si $P = NP$. Cependant, ce même problème peut être résolu en temps polynomial si nous considérons que la fonction propositionnelle Σ est une fonction de Horn.

Une question en a ainsi découlé : “Est-il possible de calculer une fonction de Horn Σ' à partir d’une fonction propositionnelle quelconque Σ tel que la plupart des inférences que l’on peut obtenir à partir de Σ puissent être obtenues en temps polynomial à partir de Σ' ?”

Cette question revient donc à savoir s’il est possible de faire une approximation d’une fonction propositionnelle quelconque Σ par une fonction de Horn Σ' et ainsi, permettre une accélération du processus de vérification des inférences. Selman et Kautz ont ainsi étudié dans [91] cette question et proposent une méthode de résolution qu’ils appellent “approximation de Horn” (traduction littérale de “Horn approximation”). Cette méthode consiste à calculer deux fonctions de Horn encadrant la fonction propositionnelle de départ. La première fonction représente ainsi une fonction de Horn minimale Σ_{lub} tel que l’ensemble des modèles de Σ soit inclus dans l’ensemble des modèles de Σ_{lub} . On l’appelle, dans la littérature, la plus petite borne supérieure de Horn ou bien l’enveloppe

de Horn. A l'inverse, la deuxième fonction représente ainsi une fonction maximale Σ_{glb} tel que l'ensemble des modèles de Σ_{glb} soit inclus dans l'ensemble des modèles de Σ . On l'appelle, dans la littérature, la plus grande borne inférieure de Horn ou le noyau de Horn. Nous pouvons voir sur la figure 1 une représentation schématique de ces approximations. Malheureusement, le calcul de ces deux fonctions d'approximation restent difficiles puisqu'il est NP-difficile dans les deux cas. Divers travaux ont ainsi été menés sur ce sujet [32, 21].

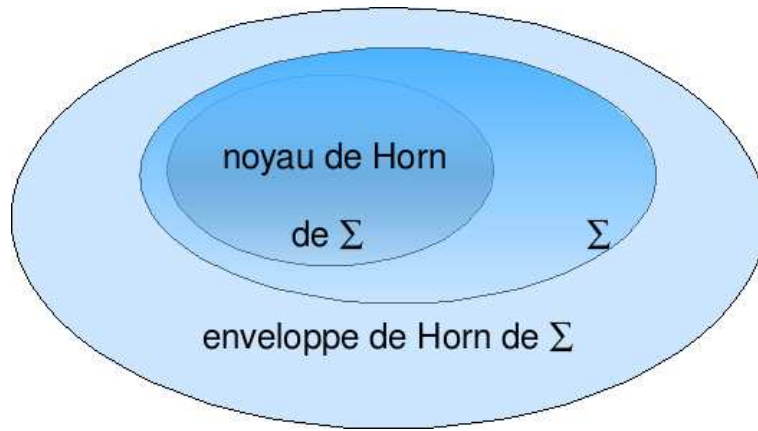


FIG. 1 – Représentation du noyau et de l'enveloppe de Horn d'une théorie Σ représentée par ses modèles

De ce fait, d'autres approches ont ainsi été étudiées en considérant des cas particuliers. Eiter et al.[41] ont ainsi étudié la méthode d'approximation utilisant les enveloppes et les noyaux de Horn sur la disjonction de théories de Horn. La nécessité de l'approximation vient du fait que la disjonction de théories de Horn n'est pas forcément de Horn. Ils ont ainsi considéré le cas où les théories de Horn sont représentées sous leur forme propositionnelle, c'est à dire sous forme de FNC de Horn. Ils ont par la même occasion considéré le cas où les théories de Horn sont représentées par leurs modèles caractéristiques. Cette approche peut ainsi avoir des applications dans la compilation de connaissance. Si nous considérons que nous avons plusieurs théories de Horn représentant des bases de connaissances, il est ainsi plus efficace de tester si une fonction φ est inférée à partir de la disjonction de ces théories plutôt que de tester cette inférence sur chacune d'entre elle.

De la même manière, ces mêmes auteurs ont étudié dans [44] ces mêmes méthodes

d'approximations sur la différence de deux théories de Horn.

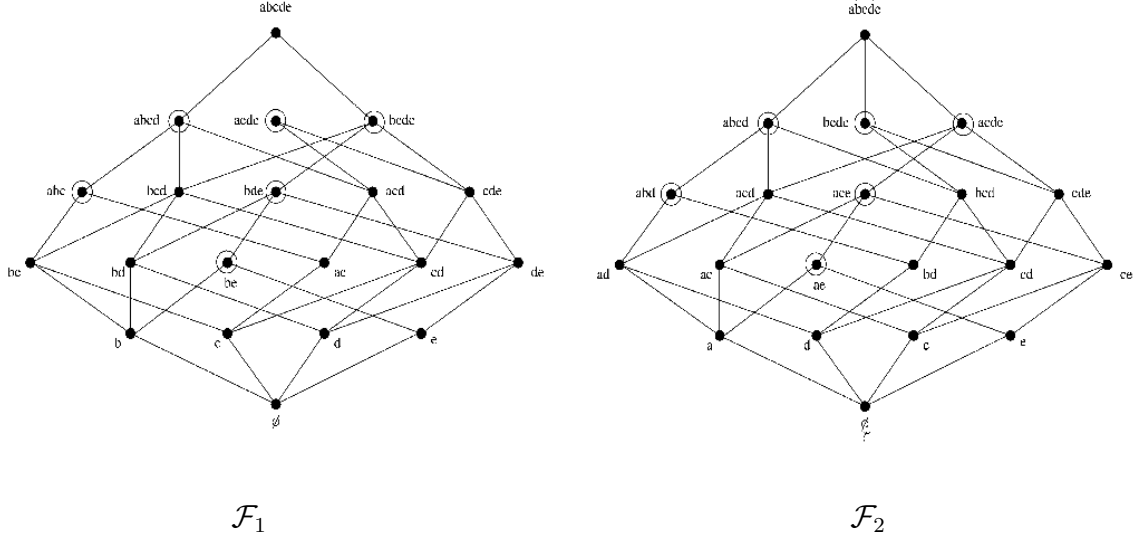
D'autre part, une étude a été menée dans [42] sur l'intersection de théorie de Horn. Ce travail a été réalisé en considérant la représentation de la théorie de Horn basée sur les modèles caractéristiques.

Nous avons vu dans la partie précédente la relation très forte qu'il existe entre les fonctions de Horn pures et les systèmes de fermeture. Ainsi, les différentes opérations évoquées précédemment sur les théories de Horn peuvent être vues comme des opérations ensemblistes sur les systèmes de fermeture. Le calcul d'une enveloppe de Horn de la disjonction de théories de Horn correspondra au calcul de la borne supérieure de systèmes de fermeture. De la même manière, le calcul de l'intersection de théories de Horn correspondra au calcul de la borne inférieure de systèmes de fermeture. Concernant la différence de deux théories de Horn, elle n'a pas d'équivalence propre dans la littérature sur les systèmes de fermeture, nous considérerons juste donc qu'elle représente la différence de deux systèmes de fermeture.

Le but de cette partie est donc d'étudier ces différentes opérations sur les systèmes de fermeture en considérant les différentes représentations de ceux-ci vus dans la partie précédente. Nous considérerons ici que le nombre de systèmes de fermeture considérés pour les opérations de borne supérieure et de borne inférieure est fini et borné par une constante k connue.

Pour illustrer ces différentes opérations relatives aux représentations considérées, nous exploiterons dans la suite de cette partie l'exemple 10. Dans cet exemple nous considérons seulement deux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 qui sont ici représentés par leur diagramme de Hasse. Les éléments entourés de chaque système de fermeture correspondent aux éléments inf-irréductibles de chacun d'entre eux. Les bases d'implications présentes sous les deux diagrammes correspondent ainsi aux bases de Guigues Duquenne $\Sigma_{\mathcal{F}_1}$ and $\Sigma_{\mathcal{F}_2}$ respectives pour \mathcal{F}_1 et \mathcal{F}_2 .

Exemple 10



$$\Sigma_{\mathcal{F}_1} = \{a \rightarrow c, ce \rightarrow d\}$$

$$\Sigma_{\mathcal{F}_2} = \{b \rightarrow d, de \rightarrow c\}$$

La suite de cette partie sera organisée de la manière suivante :

Dans le chapitre 3, nous nous intéressons à la borne inférieure de systèmes de fermeture. Nous verrons comment se caractérise cette opération selon les trois représentations que sont la base d'implications, l'ensemble des inf-irréductibles et l'opérateur de fermeture. Nous verrons en particulier qu'il est possible de définir un opérateur de fermeture pour la borne inférieure de systèmes de fermeture uniquement à partir des opérateurs de fermeture de chacun d'entre eux.

Dans le chapitre 4, nous effectuons le même travail pour la borne supérieure de systèmes de fermeture. Nous nous intéressons plus particulièrement à la représentation sous forme de base d'implications. Dans le cas général, il n'existe pas d'algorithme polynomial permettant de générer une base d'implications correspondant au système de fermeture résultant de la borne supérieure de systèmes de fermeture si ceux-ci sont représentés par des bases d'implications. Nous montrons ici qu'il existe un algorithme polynomial permettant cette génération lorsque les bases d'implications des systèmes de fermeture considérés sont directes.

Enfin, dans le chapitre 5, nous exposons les résultats sur le calcul de la fermeture de la différence de deux systèmes de fermeture selon leur représentation. Cette opération n'est pas définie sur les systèmes de fermeture. Nous retranscrivons ainsi les résultats obtenus en logique sur les systèmes de fermeture. Nous faisons par la même occasion une

conclusion générale sur l'ensemble des opérations étudiées dans cette partie ainsi qu'un bilan sur les différentes complexités de ces opérations.

Chapitre 3

Borne inférieure de systèmes de fermeture

Sommaire

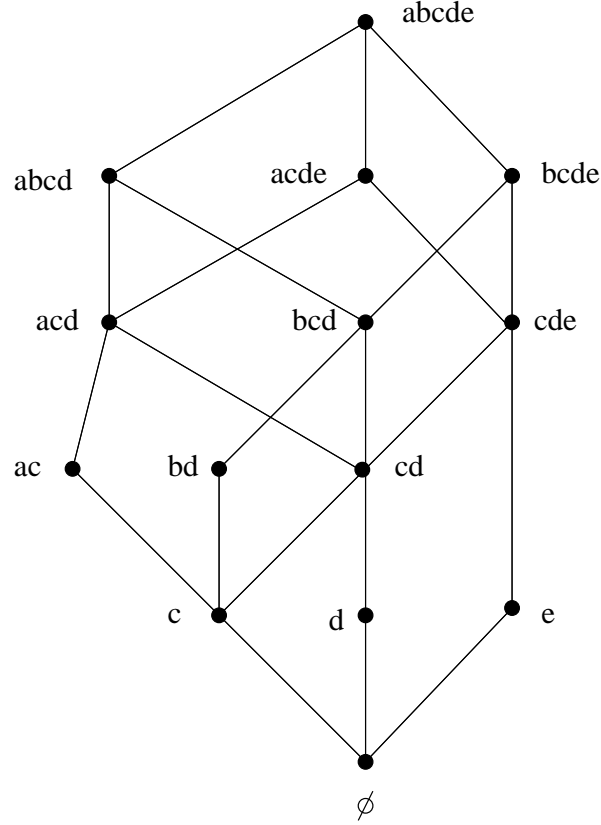
3.1	Eléments inf-irréductibles	43
3.2	Base d'implications	46
3.3	Opérateur de fermeture	48
3.4	Conclusion	49

Nous allons voir dans cette section la difficulté du calcul de la borne inférieure de systèmes de fermeture selon la représentation que nous considérons en donnée de départ. La figure 3.1 représente ici le système de fermeture obtenu par la borne inférieure des systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 de l'exemple 10.

3.1 Eléments inf-irréductibles

Nous voulons voir ici comment trouver l'ensemble des éléments inf-irréductibles du système de fermeture défini par la borne inférieure de systèmes de fermeture représentés par leurs ensembles d'éléments inf-irréductibles respectifs.

Il a été montré dans [42] que le problème de génération de l'ensemble des modèles caractéristiques de l'intersection de théories de Horn représentées par leurs modèles caractéristiques est NP-difficile. Ce problème est équivalent au problème BORNE INF :INF si nous le considérons sur les systèmes de fermeture :

FIG. 3.1 – Système de fermeture associé à $\mathcal{F}_1 \wedge \mathcal{F}_2$ de l'exemple 10.**Problème 1** :BORNE INF :INF

Données : Ensembles d'éléments inf-irréductibles $\mathcal{M}(\mathcal{F}_i)$ des systèmes de fermeture \mathcal{F}_i pour $i \in \{1, \dots, k\}$.

Résultats : L'ensemble des éléments inf-irréductibles de $\mathcal{F} = \bigwedge_{i=1}^k \mathcal{F}_i$.

L'exemple qui suit montre qu'il n'existe pas d'algorithme en temps polynomial pour le problème **BORNE INF :INF**. Cette exemple est tiré de [42] et retranscrit sur les systèmes de fermeture.

Exemple 11 Soit $G = \{x_1, \dots, x_{4n}\}$ un ensemble fini d'éléments de taille $4n$ où n est un entier.

Nous considérons quatre ensembles d'éléments de G :

V_0	V_1	V_2	V_3
$\{x_1 \dots x_n\}$	$\{x_{n+1} \dots x_{2n}\}$	$\{x_{2n+1} \dots x_{3n}\}$	$\{x_{3n+1} \dots x_{4n}\}$

Nous considérons deux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 dont les ensembles d'éléments inf-irréductibles $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$ sont définis à partir des ensembles précédents :

- $\mathcal{M}(\mathcal{F}_1) = \{G \setminus (V_2 \cup \{x_j, x_{3n+j}\}), 1 \leq j \leq n\} \cup \{G \setminus (V_2 \cup \{x_{n+j}, x_{3n+j}\}), 1 \leq j \leq n\}$
- $\mathcal{M}(\mathcal{F}_2) = \{G \setminus (V_3 \cup \{x_j, x_{2n+j}\}), 1 \leq j \leq n\} \cup \{G \setminus (V_3 \cup \{x_{n+j}, x_{2n+j}\}), 1 \leq j \leq n\}$

Notons qu'il n'y a pas d'élément inf-irréductible dans $\mathcal{M}(\mathcal{F}_1)$ qui inclut l'ensemble V_2 ni d'élément inf-irréductible dans $\mathcal{M}(\mathcal{F}_2)$ qui inclut l'ensemble V_3 .

Nous pouvons en déduire que pour tout ensemble F appartenant à $\mathcal{F}_1 \wedge \mathcal{F}_2$, F est inclus dans $V_0 \cup V_1$.

L'ensemble des éléments inf-irréductibles de $\mathcal{F}_1 \wedge \mathcal{F}_2$ est défini comme suit :

$$\mathcal{M}(\mathcal{F}_1 \wedge \mathcal{F}_2) = \{M \mid M \subseteq V_0 \cup V_1, \text{ tel que } x_j \in M \text{ ssi } x_{n+j} \notin M, 1 \leq j \leq n\}$$

Nous pouvons voir que $\mathcal{M}(\mathcal{F}_1 \wedge \mathcal{F}_2) \subseteq \mathcal{F}_1 \wedge \mathcal{F}_2$ et qu'il n'existe aucun autre élément inf-irréductible du fait qu'il n'existe pas d'ensemble fermé contenant à la fois x_j et x_{n+j} que se soit dans \mathcal{F}_1 comme dans \mathcal{F}_2 . Cette déduction est obtenue à partir de la définition de $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$.

Ainsi, la taille de $\mathcal{M}(\mathcal{F}_1 \wedge \mathcal{F}_2)$ est égale à 2^n alors que les tailles de $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$ sont égales à $2n$.

Considérons à présent le problème de décision associé au problème BORNE INF :INF.

Problème 2 : *DECISION BORNE INF :INF*

Données : Ensembles d'éléments inf-irréductibles $\mathcal{M}(\mathcal{F}_i)$ des systèmes de fermeture \mathcal{F}_i pour $i \in \{1, \dots, k\}$ et une famille d'ensemble $\mathcal{S} \subseteq \mathcal{M}(\bigwedge_{i=1}^k \mathcal{F}_i)$.

Résultats : $\mathcal{M}(\bigwedge_{i=1}^k \mathcal{F}_i) \setminus \mathcal{S} \neq \emptyset$

Théorème 6 [42] Le problème *DECISION BORNE INF :INF* est NP-complet.

Preuve. Le problème équivalent sur les théories de Horn a été démontré NP-complet [42]. ■

Du théorème, on déduit le résultat suivant :

Théorème 7 *A moins que $P = NP$, il n'existe pas d'algorithme en temps total polynomial pour le problème BORNE INF même si nous considérons seulement deux systèmes de fermeture.*

Preuve. Le problème équivalent sur les théories de Horn a été démontré NP-difficile [42]. ■

3.2 Base d'implications

Nous considérons à présent, le cas où les systèmes de fermeture sont représentés sous la forme de bases d'implications et nous voulons ainsi calculer une base d'implications pour le système de fermeture résultant de la borne inférieure de ces systèmes de fermeture.

Problème 3 BORNE INF :IMPLICATION

Données : $\Sigma_{\mathcal{F}_1}, \Sigma_{\mathcal{F}_2}, \dots, \Sigma_{\mathcal{F}_k}$ des bases d'implications pour les systèmes de fermeture $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_k$ sur un ensemble fini G .

Résultat : Σ une base d'implications de $\bigwedge_{i=1}^k \mathcal{F}_i$.

Nous montrons dans le théorème suivant que le calcul d'une base d'implications minimum correspondant à la borne inférieure des systèmes de fermeture définis par leurs bases d'implications respectives se fait en temps polynomial.

Théorème 8 Soient $\Sigma_1, \Sigma_2, \dots, \Sigma_k$ des bases d'implications minimums pour les systèmes de fermeture $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_k$. Alors $\bigcup_{i=1}^k \Sigma_i$ est une base d'implications pour $\bigwedge_{i=1}^k \mathcal{F}_i$.

Ce résultat provient du fait que les ensembles fermés de la borne inférieure des systèmes de fermeture considérés doivent satisfaire les implications des bases d'implications de chacun d'entre eux et par conséquent l'union ensembliste de chacune des base d'implications.

Il est ainsi facile de voir ici que la taille de la base d'implications pour $\bigwedge_{i=1}^k \mathcal{F}_i$ est polynomiale en la taille des Σ_i pour i dans $\{1..k\}$ sachant que nous considérons un nombre de systèmes de fermeture borné par une constante (ici k). Ainsi, du fait que le calcul d'une base d'implications minimum à partir d'une base d'implications quelconque se fait en temps polynomial (algorithme de Shock), le calcul d'une base d'implications minimum pour $\bigwedge_{i=1}^k \mathcal{F}_i$ peut donc se faire en temps polynomial.

Théorème 9 Etant données Σ_i pour i dans $\{1..k\}$ des bases d'implications pour les systèmes de fermeture \mathcal{F}_i pour i dans $\{1..k\}$ sur un ensemble G . Le calcul d'une base d'implications minimum de $\bigwedge_{i=1}^k \mathcal{F}_i$ se fait en temps polynomial.

Exemple 12 Considérons l'exemple suivant : Σ_1 et Σ_2 sont des bases d'implications minimum respectivement pour les systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 :

$$\begin{array}{ll}
\Sigma_1 = \{ & \Sigma_2 = \{ \\
& c \rightarrow d \qquad d \rightarrow c \\
& ab \rightarrow d \qquad ac \rightarrow d \\
& ad \rightarrow b \qquad ae \rightarrow b \\
& ae \rightarrow bcd \qquad bc \rightarrow d \\
& bd \rightarrow a \qquad cde \rightarrow b \\
& de \rightarrow abc \qquad \} \\
& abcd \rightarrow e \\
& \}
\end{array}$$

Ainsi, si nous calculons $\Sigma_1 \cup \Sigma_2$, nous obtenons la base d'implications suivante :

$$\begin{array}{l}
\Sigma_1 \cup \Sigma_2 = \{ \\
\qquad c \rightarrow d \\
\qquad ab \rightarrow d \\
\qquad ad \rightarrow b \\
\qquad ae \rightarrow bcd \\
\qquad bd \rightarrow a \\
\qquad de \rightarrow abc \\
\qquad abcd \rightarrow e \\
\qquad d \rightarrow c \\
\qquad ac \rightarrow d \\
\qquad ae \rightarrow b \\
\qquad bc \rightarrow d \\
\qquad cde \rightarrow b \\
\}
\end{array}$$

Nous appliquons ensuite l'algorithme de Shock afin d'obtenir une base d'implications minimum :

$$\begin{array}{l}
\Sigma_{min} = \{ \\
\qquad ac \rightarrow bde \\
\qquad bc \rightarrow ade \\
\qquad ae \rightarrow bcd \\
\qquad d \rightarrow c \\
\qquad c \rightarrow d \\
\qquad cde \rightarrow ab \\
\qquad ab \rightarrow cde \\
\}
\end{array}$$

3.3 Opérateur de fermeture

Dans cette partie, nous nous intéressons à la définition d'un opérateur de fermeture de la borne inférieure de systèmes de fermeture. Nous allons voir qu'il est possible de définir un opérateur de fermeture uniquement à partir des opérateurs de fermeture de chacun des systèmes de fermeture. Nous définissons donc l'opérateur φ dans la définition 32 et montrons que cet opérateur est un opérateur de fermeture. Nous déduisons ensuite que cet opérateur de fermeture définit bien la borne inférieure des systèmes de fermeture considérés.

Définition 32 Soient $A \subseteq G$, $\varphi_{\mathcal{F}_i}$ les opérateurs de fermeture associés aux systèmes de fermeture \mathcal{F}_i pour $i \in \{1..k\}$ et $\varphi(A) = A^0 \cup A^1 \cup A^2 \dots$

où :

- $A^0 = A$
- $A^j = A^{j-1} \cup \varphi_{\mathcal{F}_{(j \bmod k)}}(A^{j-1})$

Notons que $\varphi(A) = A^j$ quand $A^j = A^{j+t}$, pour tout t appartenant à $\{1..k-1\}$

Proposition 8 φ est un opérateur de fermeture.

Preuve. Soit $A \subseteq G$.

- $A \subseteq \varphi(A)$ par définition de $\varphi()$.
- Soit $A \subseteq D$, nous avons donc $\varphi(A) = A^0 \cup A^1 \cup A^2 \dots$ et $\varphi(D) = D^0 \cup D^1 \cup D^2 \dots$. Alors $A^j = A^{j-1} \cup \varphi_{\mathcal{F}_{(j \bmod k)}}(A^{j-1})$ et $D^j = D^{j-1} \cup \varphi_{\mathcal{F}_{(j \bmod k)}}(D^{j-1})$. Puisque $\varphi_{\mathcal{F}_i}$ sont des opérateurs de fermeture, alors $\varphi_{\mathcal{F}_i}(A) \subseteq \varphi_{\mathcal{F}_i}(D)$. Nous avons donc $\varphi(A) \subseteq \varphi(D)$.
- $\varphi(\varphi(A)) = \varphi(A^0 \cup A^1 \cup A^2 \dots) = \varphi(A^i)$. Du fait que $A^j = A^{j+t}$, pour t appartenant à $\{1..k\}$, nous avons $\varphi(A^j) = A^j = \varphi(A)$.

■

Théorème 10 Soit $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_k$ des systèmes de fermeture définis sur G . Alors $\bigcap_{i=1}^k \mathcal{F}_i = \{A \in 2^G \mid A = \varphi(A)\}$.

Preuve. Soit $F \in \mathcal{F}_1 \cap \mathcal{F}_2 \cap \dots \cap \mathcal{F}_k$. Alors $F \in \mathcal{F}_i$ pour tout i dans $\{1..k\}$. Ainsi, nous avons donc $F^j = F$ pour tout j dans $\{1..k\}$. Nous pouvons ainsi déduire à l'aide de la définition précédente que $\varphi(F) = F$. Nous en concluons que $\mathcal{F}_1 \cap \mathcal{F}_2 \cap \dots \cap \mathcal{F}_k \subseteq \{A \in 2^G \mid A = \varphi(A)\}$.

Maintenant, considérons $F \in \{A \in 2^G \mid A = \varphi(A)\}$. Si nous nous référons à la définition

précédente, nous avons $F^0 = F^j$ pour j dans $\{1..k\}$. Cela veut donc dire que $F \in \mathcal{F}_i$ pour i dans $\{1..k\}$. Nous pouvons donc conclure que $F \in \bigcap_{i=1}^k \mathcal{F}_i$ et ainsi que $\{A \in 2^G \mid A = \varphi(A)\} \subseteq \mathcal{F}_1 \cap \mathcal{F}_2 \cap \dots \cap \mathcal{F}_k$. ■

L'opérateur de fermeture que nous avons défini consiste en des itérations successives des différents opérateurs de fermeture associés à chacun des systèmes de fermeture considérés sur un ensemble fini G . Nous pouvons remarquer que le nombre d'itérations de cet opérateur est borné par le nombre de systèmes de fermeture multiplié par le nombre d'éléments de l'ensemble G . Cette borne est la conséquence directe de la condition d'extension d'un opérateur de fermeture ($A \subseteq \varphi(A)$) et de la condition d'arrêt de notre opérateur de fermeture. De part la condition d'arrêt, nous ne pouvons obtenir que k fois le même ensemble fermé par applications successives des différents opérateurs de fermeture. Si la fermeture de l'ensemble considéré est différent de l'ensemble de départ, il comporte au moins un élément de plus par la condition d'extension et est au plus égale à G .

Cet opérateur de fermeture effectue un nombre polynomial d'itérations de chaque opérateur de fermeture associé aux différents systèmes de fermeture par rapport à la taille de G et aux nombres de systèmes de fermeture pour l'obtention de la fermeture d'une ensemble.

Exemple 13 *Si nous considérons l'exemple 10, et nous voulons obtenir le fermé correspondant à l'ensemble $\{ab\}$. Nous appliquons ainsi l'opérateur de fermeture vu précédemment :*

$$\begin{aligned} ab^0 &= ab. \\ ab^1 &= abc. \\ ab^2 &= abcd \\ ab^3 &= abcd = ab^2 \end{aligned}$$

L'ensemble fermé de $\{ab\}$ est donc $\{abcd\}$ dans le système de fermeture représentant la borne inférieure de \mathcal{F}_1 et de \mathcal{F}_2 .

3.4 Conclusion

Dans ce chapitre, nous avons étudié l'opération de la borne inférieure de systèmes de fermeture selon différentes représentations de ceux-ci.

Dans un premier temps, nous voyons que si nous considérons la représentation des systèmes de fermeture par leurs ensembles d'éléments inf-irréductibles, il est difficile de

calculer l'ensemble des éléments inf-irréductibles résultant de leur borne inférieure. De manière formelle, il n'existe pas d'algorithme en temps total polynomial permettant de calculer l'ensemble des éléments inf-irréductibles hormis si $P = NP$.

Si nous considérons cette opération en prenant comme données des bases d'implications associées aux systèmes de fermeture, le calcul d'une base d'implications minimum associée au système de fermeture résultant de la borne inférieure se fait en temps polynomial. Si nous considérons le problème équivalent sur les fonctions de Horn, ce problème revient à faire une conjonction de FNC de Horn qui reste une FNC de Horn.

Enfin, nous avons défini un opérateur de fermeture associé à la borne inférieure de systèmes de fermeture. Cet opérateur de fermeture utilise uniquement les opérateurs de fermeture de chaque système de fermeture considéré. De ce fait, selon le contexte d'applications, les opérateurs de fermeture peuvent ainsi être définis par rapport aux éléments inf-irréductibles ou bien les bases d'implications ou toute autre représentation.

Chapitre 4

Borne supérieure de systèmes de fermeture

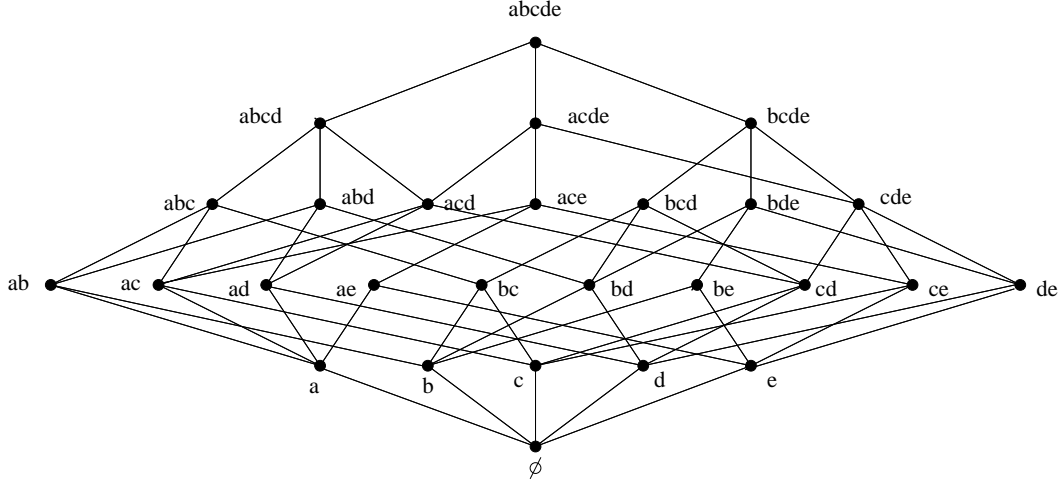
Sommaire

4.1	Eléments inf-irréductibles	51
4.2	Base d'implications	53
4.2.1	Cas particulier : bases d'implications directes	54
4.2.2	Cas général	57
4.3	Opérateur de fermeture	58
4.4	Conclusion	59

Nous allons voir à présent l'opération de borne supérieure sur les systèmes de fermeture selon la représentation que nous nous donnons de ceux-ci. Cette opération correspond à la fermeture de l'union de systèmes de fermeture. De manière équivalente sur les théories de Horn, cela correspond au calcul de l'enveloppe de Horn de la disjonction de théories de Horn. Nous pouvons voir dans la figure 4.1, le système de fermeture obtenu par la borne supérieure de \mathcal{F}_1 et \mathcal{F}_2 de l'exemple 10.

4.1 Eléments inf-irréductibles

Nous allons voir dans cette section le problème de génération des éléments inf-irréductibles de la borne supérieure de systèmes de fermeture étant donnés leurs éléments inf-irréductibles respectifs. Le problème est ainsi posé ci-dessous :

FIG. 4.1 – Système de fermeture associé à $\mathcal{F}_1 \vee \mathcal{F}_2$ de l'exemple 10**Problème 4 :BORNE SUP :INF**

Données : $\mathcal{M}(\mathcal{F}_1), \mathcal{M}(\mathcal{F}_2), \dots, \mathcal{M}(\mathcal{F}_k)$ les ensembles d'éléments inf-irréductibles des systèmes de fermeture respectifs $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_k$ sur un ensemble fini G .

Résultats : L'ensemble des éléments inf-irréductibles du système de fermeture $\mathcal{F} = \bigvee_{i=1}^k \mathcal{F}_i$.

Théorème 11 Etant donnés $\mathcal{M}(\mathcal{F}_1), \mathcal{M}(\mathcal{F}_2), \dots, \mathcal{M}(\mathcal{F}_k)$ les ensembles d'éléments inf-irréductibles des systèmes de fermeture $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_k$ sur un ensemble fini G . Le calcul de $\mathcal{M}(\bigvee_{i=1}^k \mathcal{F}_i)$ se fait en temps polynomial.

Preuve. Ce théorème est déduit directement du fait que l'union des éléments inf-irréductibles de chacun des systèmes de fermeture représentent un sur-ensemble de l'ensemble des éléments inf-irréductibles de la borne supérieure de ces systèmes de fermeture. Il suffit ainsi de supprimer les ensembles qui peuvent être obtenus par intersection d'autres ensembles pour obtenir l'ensemble voulu. ■

Exemple 14 Considérons l'exemple 10. L'union des deux ensembles d'éléments inf-irréductibles associés aux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 est le suivant :

$$\mathcal{M}(\mathcal{F}_1) \cup \mathcal{M}(\mathcal{F}_2) = \{be, abc, bde, abcd, acde, bcde, ae, ace\}$$

Dans ce cas précis, il n'existe aucun ensemble pouvant être obtenu par intersection d'autres ensembles de cette famille. Nous obtenons ainsi :

$$\mathcal{M}(\mathcal{F}_1 \vee \mathcal{F}_2) = \{be, abc, bde, abcd, acde, bcde, ae, ace\}$$

4.2 Base d'implications

Nous présentons ici le problème du calcul d'une base d'implications correspondant à la borne supérieure de systèmes de fermeture représentés par des bases d'implications. Nous décrivons ce problème ci-dessous de manière formelle :

Problème 5 BORNE SUP :IMPLICATION

Données : $\Sigma_{\mathcal{F}_1}, \Sigma_{\mathcal{F}_2}, \dots, \Sigma_{\mathcal{F}_k}$ des bases d'implications pour les systèmes de fermeture $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_k$ sur un ensemble fini G .

Résultat : Σ une base d'implications de $\bigvee_{i=1}^k \mathcal{F}_i$.

Ce problème correspond à la génération d'une enveloppe de Horn de la disjonction de fonctions de Horn représentées sous forme propositionnelle. Ce problème a été étudié dans [41]. Il en résulte qu'il n'existe pas d'algorithme polynomial pour ce problème. La proposition 9 retranscrit ce résultat sur le problème associé sur les systèmes de fermeture. L'exemple 15 est ainsi tiré de [41].

Exemple 15 [41] Soit Σ_1 et Σ_2 deux bases d'implications minimum pour les systèmes de fermeture \mathcal{F}_{Σ_1} et \mathcal{F}_{Σ_2} .

$$\begin{aligned}\Sigma_1 &= \{\emptyset \rightarrow x_0\} \\ \Sigma_2 &= \{x_1 \dots x_n \rightarrow x_0, y_1 \rightarrow x_1, \dots, y_n \rightarrow x_n\}\end{aligned}$$

La base de Guigues Duquenne obtenue pour $\mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}$ est la suivante :

$$\Sigma = \{z_1 \dots z_n \rightarrow x_0, z_i \in \{x_i, y_i\} \forall i \in \{1 \dots n\}\}$$

Proposition 9 [41] Il existe des systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 tel que $\Sigma_{\mathcal{F}_1 \vee \mathcal{F}_2}$ a une taille exponentielle en la taille de $\Sigma_{\mathcal{F}_1}$ et $\Sigma_{\mathcal{F}_2}$.

Preuve. Si nous considérons l'exemple 15, le fait qu'une base de Guigues Duquenne est minimum et que Σ possède un nombre exponentiel d'implications en la taille de Σ_1 et Σ_2 infère cette proposition. ■

Nous allons voir dans ce qui suit qu'il existe une classe particulière de base d'implications pour laquelle ce problème peut être résolu en temps polynomial. Afin d'aboutir à un tel algorithme, nous allons voir qu'il est suffisant d'avoir un algorithme permettant le calcul de la borne supérieure de deux systèmes de fermeture.

Dans la section suivante, nous considérerons la représentation des systèmes de fermeture par des bases d'implications directes.

4.2.1 Cas particulier : bases d'implications directes

Nous allons voir dans cette section que le problème BORNE SUP :IMPLICATION est polynomial en la taille de l'entrée si nous considérons que les bases d'implications prises en compte sont directes.

Proposition 10 *Le problème BORNE SUP :IMPLICATION est polynomial en la taille de l'entrée si les bases d'implications considérées sont des bases directes.*

Preuve. La preuve se déduit directement de la définition 33 et du théorème 12. ■

Définition 33 *Soit $\Sigma_1 = \{A_1 \rightarrow a_1, \dots, A_l \rightarrow a_l\}$ et $\Sigma_2 = \{B_1 \rightarrow b_1, \dots, B_m \rightarrow b_m\}$ deux bases d'implications directes associées à deux systèmes de fermeture. Nous définissons par :*

$$\Sigma_D = \{ A_i B_j \rightarrow ((A_i B_j)^{\Sigma_1} \cap (A_i B_j)^{\Sigma_2}) \setminus A_i B_j \}$$

pour tout $i \in \{1..l\}, j \in \{1..m\}$

Avec ce nouvel ensemble d'implications, nous déduisons le théorème 12 décrit comme suit :

Théorème 12 $\mathcal{F}_{\Sigma_D} = \mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}$.

Preuve. Nous devons montrer que $\mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2} \subseteq \mathcal{F}_{\Sigma_D}$.

Soit $F \in \mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}$. Deux cas sont alors possibles :

- Soit $F \in \mathcal{F}_{\Sigma_1} \cup \mathcal{F}_{\Sigma_2}$. Nous devons montrer que F est fermé par Σ_D . Soit $A \rightarrow x \in \Sigma_D$ et $A \subseteq F$. Alors $x \in (A^{\Sigma_1} \cap A^{\Sigma_2}) \setminus A$. Cela implique $x \in A^{\Sigma_1}$ et donc que $x \in F$.
- Soit $F \notin \mathcal{F}_{\Sigma_1} \cup \mathcal{F}_{\Sigma_2}$. Nous avons $F = F_1 \cap F_2$ tel que $F_1 \in \mathcal{F}_{\Sigma_1}, F_2 \in \mathcal{F}_{\Sigma_2}$. Soit $A \rightarrow x \in \Sigma_D$ et $A \subseteq F$, par construction de Σ_D , nous obtenons $x \in (A^{\Sigma_1} \cap A^{\Sigma_2}) \setminus A$. Cela implique $x \in A^{\Sigma_1}$ et $x \in A^{\Sigma_2}$ d'où $x \in F$ car $A \subseteq F_1$ et $A \subseteq F_2$.

Maintenant, il nous faut prouver que $\mathcal{F}_{\Sigma_D} \subseteq \mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}$. Ainsi, nous devons prouver que $F \in \mathcal{F}_{\Sigma_D} \setminus (\mathcal{F}_{\Sigma_1} \cup \mathcal{F}_{\Sigma_2})$ implique $F \in \mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}$ c'est à dire que $F = F^{\Sigma_1} \cap F^{\Sigma_2}$.

Supposons que $F \neq F^{\Sigma_1} \cap F^{\Sigma_2}$. Alors $\exists x \notin F$ tel que $x \in F^{\Sigma_1}$ et $x \in F^{\Sigma_2}$. Du fait que Σ_1 et Σ_2 sont des bases d'implications directes, nous avons $F^{\Sigma_1} = F \cup \bigcup \{\varphi(A) \mid A \rightarrow C \in \Sigma_1 \text{ et } A \subseteq F\}$ et $F^{\Sigma_2} = F \cup \bigcup \{\varphi(B) \mid B \rightarrow D \in \Sigma_2 \text{ et } B \subseteq F\}$. Alors il existe $A \rightarrow C \in \Sigma_1$ tel que $x \in C$ et $B \rightarrow D \in \Sigma_2$ tel que $x \in D$. Par la définition 33, $AB \rightarrow x \in \Sigma_D$ et $AB \subseteq F$. Il y a donc une contradiction puisque F est un ensemble fermé de \mathcal{F}_{Σ_D} . En conclusion,

$$\mathcal{F}_{\Sigma_D} \subseteq \mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}.$$

■

Avec le résultat du théorème 12, Σ_D représente une base d'implications pour $\mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}$. Nous pouvons donc calculer en temps polynomial une base d'implications pour $\mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}$ avec Σ_1 et Σ_2 .

Théorème 13 *Soient Σ_1, Σ_2 deux bases d'implications directes. Le calcul d'une base d'implications de $\mathcal{F}_{\Sigma_1} \vee \mathcal{F}_{\Sigma_2}$ se fait en temps polynomial.*

Exemple 16 *Considérons $\Sigma_{\mathcal{F}_1}$ et $\Sigma_{\mathcal{F}_2}$ deux bases d'implications directes respectives pour les systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 de l'exemple 10.*

$$\begin{array}{ll} \Sigma_{\mathcal{F}_1} = \{ & \Sigma_{\mathcal{F}_2} = \{ \\ & a \rightarrow c \quad b \rightarrow d \\ & ab \rightarrow c \quad ab \rightarrow d \\ & ce \rightarrow d \quad de \rightarrow c \\ & ad \rightarrow c \quad bc \rightarrow d \\ & bce \rightarrow d \quad ade \rightarrow c \\ & abd \rightarrow c \quad abc \rightarrow d \\ & ae \rightarrow cd \quad be \rightarrow cd \\ & abe \rightarrow cd \quad abe \rightarrow cd \\ & \} & \} \end{array}$$

Nous calculons $\Sigma_{\mathcal{F}_1} \vee \Sigma_{\mathcal{F}_2}$ avec l'algorithme 2. Nous avons enlevé ici toutes les implications générées à plusieurs reprises lors de l'exécution de l'algorithme.

$$\Sigma_D = \{ \begin{array}{l} ade \rightarrow c \\ abe \rightarrow cd \\ bce \rightarrow d \\ abde \rightarrow c \\ abce \rightarrow d \end{array} \}$$

Cet algorithme peut ainsi être généralisé pour un nombre constant de systèmes de fermeture. Le nombre de règles générées dans ce cadre restera polynomial par rapport aux bases d'implications prises en données puisque leur nombre est borné par une constante k . Nous présentons ainsi l'algorithme généralisé (algorithme 3) pour un nombre constant de systèmes de fermeture.

Algorithme 2 : BORNE SUP : IMPLICATION DIRECTE : 2 SF

Données : Σ_1, Σ_2 , deux bases d'implications directes respectives pour \mathcal{F}_1 et \mathcal{F}_2

Résultat : Σ_D Une base d'implications pour $\mathcal{F}_1 \vee \mathcal{F}_2$

début

$\Sigma_D = \{\}$

pour tous les implications $A \rightarrow C$ **dans** Σ_1 **faire**

pour tous les implications $B \rightarrow D$ **dans** Σ_2 **faire**

 Calculer $\text{Ext} = ((AB)^{\Sigma_1} \cap (AB)^{\Sigma_2}) \setminus AB$

si $\text{Ext} \neq \emptyset$ **alors**

$\Sigma_D = \Sigma_D \cup \{AB \rightarrow \text{Ext}\}$

 retourner Σ_D

fin

Algorithme 3 : BORNE SUP : IMPLICATION DIRECTE : k SF

Données : $\Sigma_1, \dots, \Sigma_k$, k bases d'implications directes respectives pour $\mathcal{F}_1, \dots, \mathcal{F}_k$

Résultat : Σ_D Une base d'implications pour $\bigvee_{i=1}^k \mathcal{F}_i$

début

$\Sigma_D = \{\}$

pour chaque Σ_i $i \in \{1..k\}$ **faire**

pour chaque $A_i \rightarrow B_i \in \Sigma_i$ **faire**

 . Calculer $\text{Ext} = \bigcap_{i=1}^k (A_1 \dots A_k)^{\Sigma_i} \setminus (A_1 \dots A_k)$

si $\text{Ext} \neq \emptyset$ **alors**

$\Sigma_D = \Sigma_D \cup \{A_1 \dots A_k \rightarrow \text{Ext}\}$

 retourner Σ_D

fin

4.2.2 Cas général

Nous avons vu que dans le cas général, il est impossible de calculer en temps polynomial une base d'implications de la borne supérieure de systèmes de fermeture représentés sous leur forme implicationnelle même si nous ne considérons que deux systèmes de fermeture. Kautz and Selman ont donné dans [91] un algorithme qui résout le problème équivalent dans sa formulation logique. Nous décrivons ainsi cet algorithme retranscrit sur les systèmes de fermeture.

Remarque : Kautz et Selman ont proposé cet algorithme dans le cadre de la génération d'une enveloppe de Horn associée à une fonction booléenne en forme normale conjonctive quelconque qui est un cas plus général du problème que nous étudions.

Notation 1 Soit $\Sigma_1, \Sigma_2, \dots, \Sigma_k$ des bases d'implications respectives pour $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_k$.

On note :

$$\mathcal{R} = \{X_1 X_2 \dots X_k \rightarrow x_1 \vee x_2 \vee \dots \vee x_k \mid X_i \rightarrow x_i \in \Sigma_{\mathcal{F}_i}, i \in \{1..k\}\}$$

Nous pouvons voir que toutes les règles de \mathcal{R} sont satisfaites par tous les ensembles appartenant à $\mathcal{F}_{\Sigma_1} \cup \mathcal{F}_{\Sigma_2} \cup \dots \cup \mathcal{F}_{\Sigma_k}$.

Nous décomposons ainsi \mathcal{R} en deux ensembles de règles. Le premier ensemble contient ainsi les règles d'implications tandis que le deuxième contient toutes celles qui ne sont pas des règles implicatives c'est à dire qui contiennent un "ou logique" (\vee) dans leur conclusion.

$$\begin{aligned} - \mathcal{R}_H &= \{X_1 X_2 \dots X_k \rightarrow x \mid X_i \rightarrow x_i \in \Sigma_i, i \in \{1..k\}\} \\ - \mathcal{R}_{\bar{H}} &= \{X_1 X_2 \dots X_k \rightarrow x_1 \vee \dots \vee x_k \mid X_i \rightarrow x_i \in \Sigma_i, i \in \{1..k\}\} \end{aligned}$$

Théorème 14 [91] L'algorithme 4 calcule une base d'implications pour $\mathcal{F}_1 \vee \mathcal{F}_2 \vee \dots \vee \mathcal{F}_k$.

Exemple 17 Considérons les bases de Guigues Duquenne des systèmes de fermeture de l'exemple 10 et appliquons l'algorithme 4. Nous obtenons ainsi l'ensemble de règles \mathcal{R}_H and $\mathcal{R}_{\bar{H}}$ suivants :

$$\begin{aligned} \mathcal{R}_H &= \{ade \rightarrow c, bce \rightarrow d\} \\ \mathcal{R}_{\bar{H}} &= \{ab \rightarrow c \vee d, cde \rightarrow c \vee d\} \end{aligned}$$

Après application de l'algorithme , nous obtenons :

$$\Sigma = \{ade \rightarrow c, bce \rightarrow d, abe \rightarrow c, abe \rightarrow d\}$$

Algorithme 4 : Borne supérieure :base d'implications**Données** : $\mathcal{R}_H, \mathcal{R}_{\bar{H}}$ ensemble de règles déduites de \mathcal{R} **Résultat** : Σ une base d'implications de $\mathcal{F}_1 \vee \mathcal{F}_2 \vee \dots \vee \mathcal{F}_k$ **début** **tant que** *un tel choix est possible* **faire** Essayer de trouver un couple (r_1, r_2) , $r_1 \in \mathcal{R}_H$ et $r_2 \in \mathcal{R}_{\bar{H}}$ tel que $r_1 = (xX \rightarrow a)$ et $r_2 = (Y \rightarrow \bigvee_{i=1}^l x_i \vee x)$ qui n'ai pas été considéré **si** *un tel couple existe* **alors** $r = XY \rightarrow a \vee \bigvee_{i=1}^l x_i \vee x$ **si** *r est une règle d'implication* **alors** $\mathcal{R}_H = \mathcal{R}_H \cup XY \rightarrow a$ **sinon** $\mathcal{R}_{\bar{H}} = \mathcal{R}_{\bar{H}} \cup XY \rightarrow a \vee \bigvee_{i=1}^l x_i \vee x$ $\Sigma = \mathcal{R}_H$ retourner Σ **fin**

4.3 Opérateur de fermeture

Dans cette section, le but est de définir un opérateur de fermeture correspondant à la borne supérieure de systèmes de fermeture n'ayant pour information sur ceux-ci que leurs éléments inf-irréductibles respectifs.

Nous définissons ainsi un opérateur $\phi()$ que nous démontrons être un opérateur de fermeture. Nous déduisons ensuite que $\phi()$ définit l'ensemble des éléments fermés de la borne supérieure des systèmes de fermeture considérés.

Définition 34 Soit $\phi(A) = \bigcap \{B \in \bigcup_{i=1}^k \mathcal{M}(\mathcal{F}_i) \mid A \subseteq B\}$

Proposition 11 ϕ est un opérateur de fermeture.

Preuve. Soit $A \subseteq G$.

- $A \subseteq \phi(A)$ par définition de $\phi()$.
- Soit $A \subseteq D$. Nous avons $\phi(A) = \bigcap \{B \in \bigcup_{i=1}^k \mathcal{M}(\mathcal{F}_i) \mid A \subseteq B\}$ et $\phi(D) = \bigcap \{B \in \bigcup_{i=1}^k \mathcal{M}(\mathcal{F}_i) \mid D \subseteq B\}$. Alors, $A \subseteq \phi(A)$ et $D \subseteq \phi(D)$. Nous avons donc $\phi(A) \subseteq \phi(D)$.
- $\phi(\phi(A)) = \phi(\bigcap \{B \in \bigcup_{i=1}^k \mathcal{M}(\mathcal{F}_i) \mid A \subseteq B\}) = \bigcap \{B \in \bigcup_{i=1}^k \mathcal{M}(\mathcal{F}_i) \mid A \subseteq B\} = \phi(A)$.

■

Théorème 15 Soit $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_k$ des systèmes de fermeture sur G . Alors $\bigvee_{i=1}^k \mathcal{F}_i = \{F \in 2^G \mid F = \phi(F)\}$.

Preuve. Il nous faut montrer que $\bigvee_{i=1}^k \mathcal{F}_i \subseteq \{F \in 2^G \mid F = \phi(F)\}$. Soit $F \in \bigvee_{i=1}^k \mathcal{F}_i$.

Deux cas sont alors possibles :

- Soit $F \in \bigcup_{i=1}^k \mathcal{F}_i$. Alors F appartient au moins à l'un des systèmes de fermeture. Or $\phi(F) = \bigcap \{B \in \bigcup_{i=1}^k \mathcal{M}(\mathcal{F}_i) \mid F \subseteq B\}$. Nous avons donc $\phi(F) = F$.
- Soit $F \in \bigvee_{i=1}^k \mathcal{F}_i \setminus \bigcup_{i=1}^k \mathcal{F}_i$. Alors il existe $F_i \in \mathcal{F}_i$ pour tout i dans $\{1..k\}$ tel que $F = \bigcap_{i=1}^k F_i$. Or $\phi(F) = \bigcap \{B \in \bigcup_{i=1}^k \mathcal{M}(\mathcal{F}_i) \mid F \subseteq B\}$. Nous avons donc $\phi(F) = F$.

Ainsi, nous obtenons que $\bigvee_{i=1}^k \mathcal{F}_i \subseteq \{F \in 2^G \mid F = \phi(F)\}$.

Il nous faut alors montrer que $\{F \in 2^G \mid F = \phi(F)\} \subseteq \bigvee_{i=1}^k \mathcal{F}_i$.

Soit $F \in 2^G$ tel que $F = \phi(F)$. Alors $\phi(F) = \bigcap \{B \in \bigcup_{i=1}^k \mathcal{M}(\mathcal{F}_i) \mid F \subseteq B\}$. Nous concluons donc que $F \in \bigvee_{i=1}^k \mathcal{F}_i$ et donc que $\bigvee_{i=1}^k \mathcal{F}_i \subseteq \{F \in 2^G \mid F = \phi(F)\}$. ■

Exemple 18 Si nous considérons l'exemple 10, nous voulons obtenir l'ensemble fermé associé à l'ensemble $\{ab\}$ dans $\mathcal{F}_1 \vee \mathcal{F}_2$. Nous appliquons alors l'opérateur de fermeture obtenu sur cet ensemble :

$$\phi(ab) = \bigcap \{B \in \mathcal{M}(\mathcal{F}_1) \cup \mathcal{M}(\mathcal{F}_2) \mid ab \subseteq B\} = abcd \cap abc \cap abd = ab.$$

L'opérateur de fermeture défini pour la borne supérieure de systèmes de fermeture considère l'union des ensembles d'éléments inf-irréductibles de chaque système de fermeture. Le calcul de la fermeture d'un ensemble donné se fait donc en temps polynomial par rapport aux ensembles des éléments inf-irréductibles de chaque système de fermeture considéré.

4.4 Conclusion

Nous avons vu pour l'opération de borne inférieure sur les systèmes de fermeture que la difficulté de calcul diffère selon la représentation prise en compte. Si nous considérons la représentation par les ensembles d'éléments inf-irréductibles, l'opération de borne inférieure est une opération difficile. A l'inverse, si nous considérons l'opération de borne supérieure sur cette même représentation, il existe un algorithme polynomial en la taille

de l'entrée qui permet de la réaliser.

De manière duale, si nous considérons la représentation des systèmes de fermeture par les bases d'implications, l'opération de borne inférieure se fait en temps polynomial alors que l'opération de borne supérieure peut engendrer, dans le cas général, une explosion combinatoire. Nous nous sommes ainsi intéressés à une classe particulière de bases d'implications que sont les bases d'implications directes. Nous proposons, dans ce cas précis, un algorithme en temps polynomial permettant de résoudre l'opération de borne supérieure de systèmes de fermeture représentés sous forme implicationnelle.

Le problème de génération d'une base d'implications de la borne supérieure de systèmes de fermeture représentés sous forme implicationnelle a été récemment classé [45]. Il a ainsi été montré qu'il n'existe pas d'algorithme polynomial pour résoudre ce problème en temps total polynomial hormis si $P = NP$.

Il existe des algorithmes permettant de générer une base d'implications directe à partir d'une base d'implications quelconque. Ces algorithmes restent exponentiels en la taille de l'entrée et de la sortie mais ont de bonnes propriétés en pratique.

Nous pouvons ainsi adapter l'algorithme sur les bases d'implications directes au cas général. Il suffit pour cela d'appliquer un algorithme de génération de bases d'implications directes à partir de bases d'implications quelconques et d'appliquer l'algorithme 2 pour répondre au problème 5. Il serait ainsi intéressant de tester, de manière pratique, l'algorithme ainsi obtenu est de comparer ses performances avec les autres algorithmes existants.

Chapitre 5

Différence de deux systèmes de fermeture

Sommaire

5.1	Eléments inf-irréductibles	61
5.2	Base d'implications	64
5.3	Opérateur de fermeture	65
5.4	Conclusion	66

Nous allons voir dans cette section l'opération de différence de deux systèmes de fermeture selon la représentation que nous considérons en donnée de départ. Cette opération n'est pas définie formellement sur les systèmes de fermeture. La différence entre deux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 , notée $\mathcal{F}_1 \setminus \mathcal{F}_2$, consiste à construire la famille des ensembles appartenant à \mathcal{F}_1 qui n'appartiennent pas à \mathcal{F}_2 . Il est clair ici que la différence de deux systèmes de fermeture n'est pas toujours un système de fermeture. Nous considérerons donc ici le plus petit système de fermeture qui contient les ensembles de la différence des deux systèmes de fermeture. Ce problème correspond au calcul de l'enveloppe de Horn de la différence de deux théories de Horn en logique. La figure 5 montre ainsi le système de fermeture obtenu en faisant la différence des systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 de l'exemple 10.

5.1 Eléments inf-irréductibles

Nous voulons voir ici comment calculer les éléments inf-irréductibles de la fermeture de la différence de deux systèmes de fermeture à partir de leurs ensembles d'éléments inf-irréductibles respectifs.

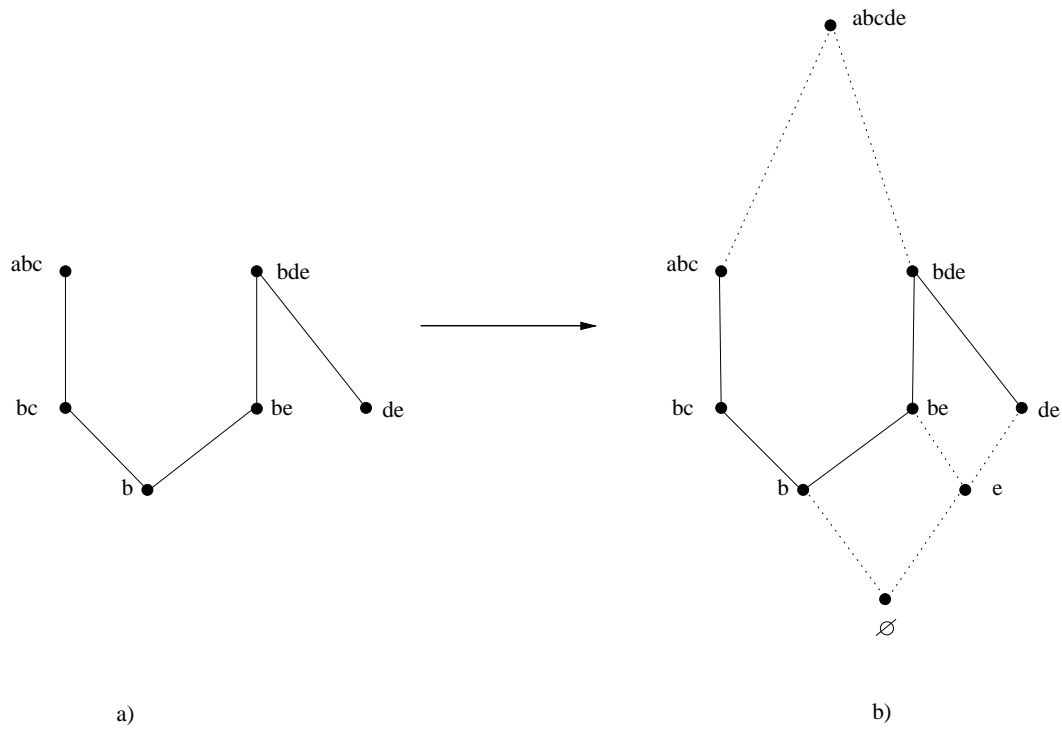


FIG. 5.1 – a) Famille d'ensembles $\mathcal{F}_1 \setminus \mathcal{F}_2$, b) Système de fermeture associé à la fermeture de $\mathcal{F}_1 \setminus \mathcal{F}_2$.

Problème 6 :*DIFFERENCE* :*INF*

Données : Ensembles d'éléments inf-irréductibles $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$ des système de fermeture \mathcal{F}_1 et \mathcal{F}_2 .

Résultats : L'ensemble des éléments inf-irréductibles de la fermeture de $\mathcal{F} = \mathcal{F}_1 \setminus \mathcal{F}_2$.

Le problème 6 a été étudié dans sa formulation logique dans [44]. Les auteurs ont ainsi montré que ce problème peut être résolu en temps polynomial. Nous retranscrivons ainsi ce résultat sous sa formulation basée sur les systèmes de fermeture.

Définition 35 Soit $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$ les ensembles d'éléments inf-irréductibles respectivement aux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 . On définit :

$$\mathcal{S} = \mathcal{M}(\mathcal{F}_1) \cup \{M \cap M' \mid M, M' \in \mathcal{M}(\mathcal{F}_1)\}$$

et

$$\begin{aligned}\mathcal{S}_1 &= \{S \in \mathcal{S} \mid S \notin \mathcal{F}_2\} \\ \mathcal{S}_2 &= \mathcal{S} \setminus \mathcal{S}_1\end{aligned}$$

Théorème 16 [44] Soient $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$ les ensembles d'éléments inf-irréductibles respectivement aux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 . Alors \mathcal{S}_1 est un sur-ensemble des inf-irréductibles de la fermeture de la différence de \mathcal{F}_1 et \mathcal{F}_2 .

Corollaire 2 [44] Soient $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$ les ensembles d'éléments inf-irréductibles respectivement aux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 . Alors l'ensemble des inf-irréductibles de la fermeture de la différence de \mathcal{F}_1 et \mathcal{F}_2 peut être calculé en temps polynomial en la taille de $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$.

Preuve. Les preuves de ces théorèmes sont donnés dans [44]. ■

L'algorithme 5 présente la méthode pour calculer l'ensemble des éléments inf-irréductibles de la fermeture de la différence de deux systèmes de fermeture étant donnés leurs éléments inf-irréductibles respectifs. La fonction Réduire() supprime les ensembles pouvant être obtenus par intersection d'autres ensembles.

Exemple 19 Considérons l'exemple 10. En appliquant l'algorithme 5 nous obtenons les ensembles \mathcal{S} et \mathcal{S}_1 suivants :

$$\begin{aligned}\mathcal{S} &= \{abcd, acde, bcde, abc, bde, be, bcd, acd, cde, bc, ac, b, e\} \\ \mathcal{S}_1 &= \{abc, bde, be, bc, b\}\end{aligned}$$

L'application de la procédure **Réduire()** supprime ici l'élément b de \mathcal{S}_1 puisque $b = \{bc\} \cap \{be\}$. Nous obtenons ainsi l'ensemble d'éléments inf-irréductibles suivants pour $\mathcal{F}_1 \setminus \mathcal{F}_2$:

$$\mathcal{M}(\mathcal{F}_1 \setminus \mathcal{F}_2) = \{abc, bde, be, bc\}.$$

Algorithme 5 : Différence : représentation par les inf-irréductibles

Données : $\mathcal{M}(\mathcal{F}_1)$ et $\mathcal{M}(\mathcal{F}_2)$ les ensembles d'éléments inf-irréductibles de \mathcal{F}_1 et \mathcal{F}_2

Résultat : $\mathcal{M}(\mathcal{F}_1 \setminus \mathcal{F}_2)$ l'ensemble des éléments inf-irréductibles de $\mathcal{F}_1 \setminus \mathcal{F}_2$

.

début

$\mathcal{S} = \mathcal{M}(\mathcal{F}_1) \cup \{M \cap M' \mid M, M' \in \mathcal{M}(\mathcal{F}_1)\}$ $\mathcal{S}_1 = \{S \in \mathcal{S} \mid S \notin \mathcal{F}_2\}$ $\mathcal{M}(\mathcal{F}_1 \setminus \mathcal{F}_2) = \text{Réduire}(\mathcal{S}_1)$ retourner $\mathcal{M}(\mathcal{F}_1 \setminus \mathcal{F}_2)$
--

fin

5.2 Base d'implications

Dans cette section, nous voyons comment se traduit la différence de deux systèmes de fermeture si nous considérons leurs représentations sous forme de bases d'implications.

Problème 7 *DIFFERENCE :IMPLICATION*

Données : Σ_1, Σ_2 deux bases d'implications respectives par rapport aux système de fermeture \mathcal{F}_1 et \mathcal{F}_2 .

Résultats : Σ une base d'implications pour la fermeture de la différence entre \mathcal{F}_1 et \mathcal{F}_2 .

Il a été montré dans [44], que le problème équivalent en logique c'est à dire le calcul de l'enveloppe de Horn de la différence de deux théories de Horn, est co-NP difficile. Par ce résultat, il parait donc évident qu'il y a peu de chance qu'il existe un algorithme polynomial permettant de résoudre ce problème.

Ceci est dû au fait que la taille de la nouvelle base peut être exponentielle en la taille des deux bases d'implications prises en entrée. De plus, le problème de décision associé est lui même co-NP-complet.

Problème 8 *DECISION DIFFERENCE :IMPLICATION*

Données : Σ_1, Σ_2 , des bases d'implications respectives par rapport aux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 . Σ une base d'implications

Résultats : $\mathcal{F}_\Sigma = \text{fermeture de } \mathcal{F}_{\Sigma_1} \setminus \mathcal{F}_{\Sigma_2}$.

Théorème 17 [44] *Etant données Σ_1, Σ_2 , des bases d'implications respectives par rapport aux systèmes de fermeture \mathcal{F}_1 et \mathcal{F}_2 et Σ une base d'implications, décider si Σ représente une base d'implications pour la fermeture de la différence de \mathcal{F}_1 et \mathcal{F}_2 est co-NP-complet.*

Le théorème suivant en est ainsi déduit :

Théorème 18 [44] *Etant donné Σ_1, Σ_2 deux bases d'implications, il n'existe pas d'algorithme en temps total polynomial permettant de calculer une base d'implications de la fermeture de la différence des systèmes de fermeture hormis si $P = NP$.*

Ce théorème a été prouvé dans [44] sur le problème équivalent en logique propositionnelle pour l'enveloppe de Horn de la différence de deux théories de Horn.

5.3 Opérateur de fermeture

Nous allons voir à présent l'opérateur de fermeture que l'on peut définir à partir des éléments inf-irréductibles des deux systèmes de fermeture pris en considération pour la fermeture de la différence.

Nous définissons ainsi un opérateur ψ et nous démontrons qu'il représente un opérateur de fermeture. Nous montrons ensuite l'équivalence entre la famille d'ensembles fermés par ψ et les ensembles fermés de la différence des deux systèmes de fermeture pris en compte. Pour ceci nous nous servons des ensembles \mathcal{S} et \mathcal{S}_1 définis dans la section 5.1 concernant la représentation par les éléments inf-irréductibles. L'opérateur de fermeture ainsi défini permet un calcul en temps polynomial de la fermeture d'un ensemble donné par rapport aux éléments inf-irréductibles de chaque système de fermeture.

Définition 36 Soit $\psi(A) = \bigcap \{B \in \mathcal{S}_1 \mid A \subseteq B\}$.

Proposition 12 $\psi()$ est un opérateur de fermeture.

Preuve. Soit $A \subseteq G$.

- $A \subseteq \psi(A)$ par définition de $\psi()$.
- Soit $A \subseteq D$. Nous avons $\psi(A) = \bigcap \{B \in \mathcal{S}_1 \mid A \subseteq B\}$ et $\psi(D) = \bigcap \{B \in \mathcal{S}_1 \mid D \subseteq B\}$. Cela implique que $A \subseteq \psi(A)$ et $D \subseteq \psi(D)$. Nous avons donc $\psi(A) \subseteq \psi(D)$.
- $\psi(\psi(A)) = \psi(\bigcap \{B \in \mathcal{S}_1 \mid A \subseteq B\}) = \bigcap \{B \in \mathcal{S}_1 \mid A \subseteq B\} = \psi(A)$.

■

Théorème 19 Soient \mathcal{F}_1 et \mathcal{F}_2 deux systèmes de fermeture sur G . Alors $\mathcal{F}_1 \setminus \mathcal{F}_2 = \{F \in 2^G \mid F = \psi(F)\}$.

Preuve. Nous savons par définition que $\mathcal{F}_1 \setminus \mathcal{F}_2 = \{F \in G \mid F = \bigcap \{M \in \mathcal{M}(\mathcal{F}_1 \setminus \mathcal{F}_2) \mid F \subseteq M\}$.

D'après le théorème 16, nous avons \mathcal{S}_1 est un sur-ensemble des éléments inf-irréductibles de $\mathcal{F}_1 \setminus \mathcal{F}_2$.

Nous avons donc $\mathcal{F}_1 \setminus \mathcal{F}_2 = \{F \in 2^G \mid F = \psi(F)\}$. ■

Exemple 20 *Considérons l'exemple 10. Nous voulons obtenir l'ensemble fermé de $\{b\}$ dans $\mathcal{F}_1 \setminus \mathcal{F}_2$. Nous appliquons alors l'opérateur de fermeture obtenu sur cet ensemble :*

$$\psi(b) = \bigcap \{B \in \mathcal{S}_1 \mid \{b\} \subseteq B\} = \{be\} \cap \{bc\} \cap \{b\} = \{b\}.$$

5.4 Conclusion

L'opération de différence entre deux systèmes de fermeture est la dernière opération étudiée dans cette partie. Nous voyons ici, de la même manière que les opérations précédentes, que les difficultés de calcul ne sont pas les mêmes selon les représentations que nous considérons des systèmes de fermeture.

Tout d'abord, nous voyons que le calcul de l'ensemble des éléments inf-irréductibles de la fermeture de la différence de deux systèmes de fermeture s'effectue en temps polynomial à partir des éléments inf-irréductibles de chacun d'entre eux.

En revanche, cette même opération s'avère difficile si nous considérons une représentation implicationnelle des systèmes de fermeture. Plus formellement, il n'existe pas d'algorithme en temps total polynomial permettant de calculer une base d'implications de la fermeture de la différence des systèmes de fermeture hormis si $P = NP$.

Si nous regardons à présent, de manière globale, l'ensemble des opérations de cette partie, nous remarquons que les difficultés de calcul ne sont pas liées à la représentation considérée des systèmes de fermeture. Si nous considérons l'opération de borne inférieure, nous voyons que pour la représentation par les éléments inf-irréductibles, l'opération est difficile tandis que si nous considérons la représentation implicationnelle, l'opération se fait en temps polynomial. A l'inverse, pour les opérations de borne supérieure et de différence, la difficulté porte sur le calcul sur la représentation implicationnelle tandis que pour la représentation par les éléments inf-irréductibles, cette opération se fait en temps polynomial.

Ces opérations seront ainsi plus ou moins applicables selon la représentation des

	Borne inférieure	Borne supérieure	Différence
Inf-irréductibles	NP-difficile	Polynomial	Polynomial
Bases d'implications	Polynomial	co-NP-difficile	co-NP-difficile
Bases d'implications directes	Polynomial	Polynomial	ouvert
Opérateur de fermeture	Polynomial	Polynomial	Polynomial

TAB. 5.1 – Récapitulatif des complexités des différentes opérations sur les systèmes de fermeture selon leur représentation.

données que nous considérons. Les complexités de chaque opération par rapport à leurs différentes représentations sont exposées dans la table 5.1

Quatrième partie

Problème de génération des bases d'implications mixtes

Introduction

Nous avons vu dans la partie précédente qu’une base d’implications est une représentation compacte d’un ensemble de données. La génération de telles bases a été très étudiée et reste toujours un problème d’actualité. Cette étude est ainsi présente dans de nombreux domaines de l’informatique tel que le data mining, la théorie des graphes, la logique... Ainsi le problème d’extraction de règles d’association à partir d’un ensemble de données [2] est un des domaines où cette recherche est toujours d’actualité. Une règle d’association est une implication de la forme $X \rightarrow Y$ [sup,conf], où X et Y sont des sous-ensembles d’attributs (appelés itemsets) tel que $X \cap Y = \emptyset$ et tel que conf représente la confiance et sup représente le support. Le support d’une règle est défini par la probabilité qu’un ensemble d’objets possède X et Y tandis que la confiance représente la probabilité que Y appartienne à un objet sachant que X appartient à cet objet.

Une approche intéressante fut l’utilisation de l’analyse de concepts formels afin de calculer un ensemble réduit de règles d’associations. De ce fait, plusieurs études ont conduit à la génération de représentations concises de ces ensembles de règles telles que la base des générateurs minimaux (ou base générique) [85] et la base de Guigues-Duquenne [38, 54]. Les deux dernières bases de règles d’associations ont la propriété d’être composées uniquement de règles dont la confiance est égale à 1 que l’on nomme règles exactes. Elles correspondent ainsi à des bases d’implications.

La plupart de ces études ont été développées afin de trouver des associations positives entre les différents éléments considérés. Si nous prenons l’exemple bien connu du panier de la ménagère, une association positive possible serait : *“Les consommateurs qui achètent de la bière et du whisky, achètent aussi des cacahuètes”*.

D’autres règles peuvent être intéressantes si nous considérons la négation sur les éléments. Nous appellerons ici ces règles “règles mixtes”. Cette négation peut ainsi permettre d’obtenir de nouvelles informations pertinentes que ne permettent pas les règles d’associations positives. Si nous reprenons l’exemple du panier de la ménagère, une association avec négation pourra être du type *“les consommateurs qui achètent de la bière et du whisky, achètent aussi des cacahuètes mais n’achètent pas de couche culotte”*.

L'ajout dans l'implication du non achat des couches culottes rajoute une information qui permet de faire de nouvelles déductions et d'étayer les analyses faites sur les données considérées.

Différents travaux ont ainsi été menés sur ce sujet. La notion d'association négative entre itemsets fut étudiée en premier par Brin et al. [20]. Le manque d'expressivité des règles d'associations positives conduit ainsi à la recherche de nouvelles informations que peut amener les règles d'association mixtes. L'introduction de la négation dans les règles d'associations peut ainsi être utile pour l'identification d'exceptions [93] et de corrélations négatives [3, 90]. Elle a permis d'aboutir à différentes perspectives dans la littérature et se révèle être un défi important dans les différents domaines évoqués. Différentes techniques et mesures de qualité sont explorées dans divers travaux de recherche récents [19, 102].

Malheureusement, il n'existe que très peu d'algorithmes de génération efficaces permettant d'obtenir un ensemble valide et complet de règles d'association mixtes à partir d'un ensemble de données (La validité d'un ensemble de règles d'association représente le fait que toutes les règles d'association satisfont l'ensemble de données et la complétude d'un ensemble de règles d'association le fait que toutes les règles d'association qui satisfont un ensemble de données peuvent être déduites de l'ensemble de règles d'association généré). Dans cette partie, nous nous intéressons donc au problème de génération d'une base d'implications mixtes (c'est à dire prenant en compte la présence et l'absence des attributs dans le contexte) à partir d'un contexte formel $R = (G, M, I)$. Ce problème est ainsi équivalent à la génération de règles d'association dont la confiance est égale à 1.

Nous verrons ainsi les différentes problématiques qui peuvent exister sur ces bases d'implications selon le choix de données de départ que nous considérons. Nous regarderons ainsi la faisabilité et la complexité de chacune.

Nous nous focaliserons plus particulièrement sur le problème de la génération d'une base d'implications mixtes à partir des bases d'implications purement positives et purement négatives associées à un contexte R .

Cette partie est structurée comme suit :

Le chapitre 6 présente les notations et les définitions que nous allons utiliser pour les bases d'implications mixtes. Nous introduisons une méthode naïve de génération des règles d'implications à partir d'un contexte formel $R = (G, M, I)$. Nous regardons ensuite différentes problématiques liées à la génération des bases d'implications purement positives, purement négatives et mixtes. Nous verrons que la plupart des problèmes évoqués sont à ce jour encore ouverts.

Dans le chapitre 7, nous nous concentrons plus particulièrement sur le problème de la génération d'une base d'implications mixtes à partir des bases d'implications génériques purement positives et purement négatives associées à un contexte R . Nous verrons ainsi, qu'à partir de cette information, nous n'avons aucune garantie d'obtenir une base d'implications mixtes complète pour un contexte R considéré. Cependant, nous exhiberons un certain nombre de propriétés et de règles d'inférence qui nous permettent de déduire un certain nombre d'implications mixtes de R .

Nous montrerons aussi que dans le cas particulier d'un contexte réduit, nous obtenons une nouvelle règle d'inférence qui nous permet de montrer la validité et la complétude de la base d'implications mixtes déduite pour un contexte à partir des règles d'inférences que nous avons trouvées en considérant en données les bases d'implications génériques positives et négatives de ce contexte. Un contexte particulier, que l'on nomme "contexte complet" sera traité indépendamment pour arriver à ce résultat.

Chapitre 6

Définitions et Problématique

Sommaire

6.1	Définitions et notations	75
6.1.1	Contexte complémentaire et apposition de contextes	75
6.1.2	Implications positives, négatives et mixtes	76
6.2	Problématique	78

Nous allons voir dans ce chapitre les définitions et les notations qui vont nous servir pour les implications mixtes tout au long de cette partie. Nous verrons ensuite les problématiques qui peuvent se poser sur ce sujet ainsi que l'analyse de leur complexité connue à ce jour.

6.1 Définitions et notations

6.1.1 Contexte complémentaire et apposition de contextes

Définition 37 *Contexte complémentaire :*

Soit $R = (G, M, I)$ un contexte formel. On définit par $\tilde{R} = (G, \tilde{M}, (M \times G) \setminus I)$ le contexte complémentaire de R .

Le contexte formel complémentaire \tilde{R} a pour but de représenter l'absence des attributs dans les objets du contexte R . Il matérialise l'aspect négatif du contexte de départ. Ainsi, si nous considérons une représentation matricielle du contexte complémentaire \tilde{R} les 1 correspondront aux 0 dans le contexte R et vice et versa.

Notation 2 Soit $R = (G, M, I)$ un contexte formel. On note par :

- \tilde{M} : l'ensemble des attributs négatifs associés à M .

$R \tilde{R}$	a	b	c	d	\tilde{a}	\tilde{b}	\tilde{c}	\tilde{d}
1	1	0	0	1	0	1	1	0
2	0	1	0	1	1	0	1	0
3	1	1	0	0	0	0	1	1
4	0	1	1	1	1	0	0	0

TAB. 6.1 – L'apposition de deux contextes complémentaires $R|\tilde{R}$.

- \tilde{x} : la négation d'un élément de M i.e un élément de \tilde{M} .
- \tilde{X} : un ensemble d'éléments inclus dans \tilde{M} .

Définition 38 *Apposition de contextes :*

Soit $R_1 = (G, M_1, I_1)$ et $R_2 = (G, M_2, I_2)$ deux contextes formels. On définit par $R = R_1|R_2 = (G, M_1 \dot{\cup} M_2, I_1 \dot{\cup} I_2)$ l'apposition de R_1 et R_2 .

Si nous considérons la représentation matricielle des contextes formels, l'apposition de deux contextes R_1 et R_2 correspond à la concaténation des matrices partageant le même ensemble d'objets [54]. Nous pouvons ainsi voir dans la table 6.1 l'apposition de deux contextes formels complémentaires.

6.1.2 Implications positives, négatives et mixtes

Nous considérons ici les implications sur l'ensemble des attributs d'un contexte formel. Nous associerons donc à chaque contexte son système de fermeture sur l'ensemble des attributs de ce contexte.

Notation 3 Soit $R = (G, M, I)$ un contexte formel. on notera par \mathcal{F}_R le système de fermeture sur l'ensemble M des attributs de R .

Nous allons ici distinguer trois types d'implications.

- Implications positives : $X \rightarrow Y$ avec $X \subseteq M$ et $Y \subseteq M$.
- Implications négatives : $\tilde{X} \rightarrow \tilde{Y}$ avec $\tilde{X} \subseteq \tilde{M}$ et $\tilde{Y} \subseteq \tilde{M}$.
- Implications mixtes : $X\tilde{Z} \rightarrow Y\tilde{T}$ avec $\tilde{Z}, \tilde{T} \subseteq \tilde{M}$ et $X, Y \subseteq M$.

Nous pouvons remarquer que les implications positives et négatives sont des cas particuliers des implications mixtes. L'ensemble des implications mixtes associé à un contexte formel représente ainsi un sur-ensemble des implications positives habituellement considérées.

positives	négatives	mixtes	
$c \rightarrow bd$	$\tilde{b} \rightarrow \tilde{c}$	$a \rightarrow \tilde{c}$	$abd \rightarrow \tilde{a}\tilde{b}\tilde{c}\tilde{d}$
$ac \rightarrow bd$	$\tilde{d} \rightarrow \tilde{c}$	$bd \rightarrow \tilde{a}$	$ac \rightarrow \tilde{a}\tilde{b}\tilde{c}\tilde{d}$
$abd \rightarrow c$	$\tilde{a}\tilde{b} \rightarrow \tilde{c}\tilde{d}$	$ad \rightarrow \tilde{b}\tilde{c}$	$\tilde{a}\tilde{d} \rightarrow \tilde{a}\tilde{b}\tilde{c}\tilde{d}$
	$\tilde{a}\tilde{d} \rightarrow \tilde{b}\tilde{c}$	$c \rightarrow \tilde{a}bd$	$\tilde{b}\tilde{d} \rightarrow \tilde{a}\tilde{a}\tilde{b}\tilde{c}\tilde{d}$
	$\tilde{b}\tilde{d} \rightarrow \tilde{a}\tilde{c}$	$ab \rightarrow \tilde{c}\tilde{d}$	$\tilde{a}\tilde{b} \rightarrow \tilde{a}\tilde{b}\tilde{c}\tilde{d}$
		$\tilde{a} \rightarrow bd$	$a\tilde{a} \rightarrow \tilde{b}\tilde{b}\tilde{c}\tilde{d}$
		$\tilde{b} \rightarrow ad\tilde{c}$	$b\tilde{b} \rightarrow \tilde{a}\tilde{a}\tilde{c}\tilde{d}$
		$\tilde{d} \rightarrow ab\tilde{c}$	$c\tilde{c} \rightarrow \tilde{a}\tilde{a}\tilde{b}\tilde{d}$
		$\tilde{a}\tilde{c} \rightarrow bd$	$d\tilde{d} \rightarrow \tilde{a}\tilde{a}\tilde{b}\tilde{b}\tilde{c}$
		$b\tilde{c}d \rightarrow \tilde{a}$	$\tilde{b}\tilde{c} \rightarrow \tilde{a}\tilde{a}\tilde{b}\tilde{c}\tilde{d}$
			$c\tilde{d} \rightarrow \tilde{a}\tilde{a}\tilde{b}\tilde{b}\tilde{c}\tilde{d}$

TAB. 6.2 – Implications positives, négatives et mixtes associées au contexte R de la table 6.1.

Une solution simple pour le problème de génération des règles mixtes d'un contexte formel donné R consiste à faire l'apposition de R avec \tilde{R} . Le calcul des implications mixtes de la relation R revient ainsi à calculer les implications positives de la relation $R|\tilde{R}$. Malheureusement, cette méthode n'est pas très efficace. Un des facteurs majeurs est le fait que, dans des applications réelles, les collections de données représentées sont relativement éparées et ainsi les contextes formels complémentaires correspondants sont denses. Ainsi la génération d'une base d'implications mixtes peut être très coûteuse du fait du doublement du nombre d'attributs du contexte formel et du nombre de concepts associés à ce nouveau contexte.

Notation 4 Soit $R = (G, M, I)$ un contexte formel. On notera par :

- Σ_R : une base d'implications pour le contexte R (c'est à dire les implications positives de R).
- $\Sigma_{\tilde{R}}$: une base d'implications pour le contexte \tilde{R} (c'est à dire les implications négatives de R).
- $\Sigma_{R|\tilde{R}}$: une base d'implications pour le contexte $R|\tilde{R}$ (c'est à dire les implications mixtes de R).

Nous pouvons voir dans la table 6.2 les bases d'implications génériques positives, négatives et mixtes du contexte formel R de la table 6.1.

Pour chaque ensemble d'implications, nous distinguons deux types d'implications :

- Les implications à support strictement positif du type $X \rightarrow Y$: Ce sont les implications tel qu'au moins un objet vérifie cette implication.
- Les implications à support nul du type $X \rightarrow M [0]$: Ce sont les implications telles qu'aucun objet de la relation comporte tous les attributs de la prémisse. Ils représentent, pour les systèmes de fermeture, les ensembles non fermés tels que leur fermeture est égale à M .

Ces notions seront appliquées par la suite sur les différents types d'implications c'est à dire les implications positives, négatives et mixtes.

6.2 Problématique

Dans cette section, nous identifions un ensemble de problèmes relatifs à la génération de bases d'implications positives, négatives et mixtes à partir d'un contexte formel $R = (G, M, I)$ donné ou non. Nous donnons pour chacun d'entre eux leur complexité associée connue à ce jour.

Problème 9 *GENERATION IMPLICATIONS POSITIVES (GIP)*

Données : *Un contexte formel $R = (G, M, I)$.*

Résultat : Σ_R *une base d'implications associée à R .*

Le problème GIP a pour but de calculer une base d'implications étant donné un contexte formel R . Ce problème est un problème classique en analyse formelle de concepts. L'étude intensive de ce problème dans différents domaines a conduit à l'apparition de divers algorithmes pour la génération de ces bases d'implications [38, 54, 85]. La complexité de ce problème est à ce jour toujours ouvert même si nous considérons des cas particuliers de celui-ci [47].

Problème 10 *GENERATION IMPLICATIONS NEGATIVES (GIN)*

Données : *Un contexte formel $R = (G, M, I)$.*

Résultat : $\Sigma_{\bar{R}}$.

Le problème GIN consiste à générer un ensemble d'implications négatives étant donné un contexte formel R . Ce problème est ainsi équivalent au problème GIP à une étape de prétraitement près. Pour obtenir l'ensemble des implications négatives, il suffit de calculer

le contexte complémentaire \tilde{R} du contexte R et d'appliquer les algorithmes de génération proposés pour le problème GIP sur ce nouveau contexte ainsi obtenu.

Problème 11 *GENERATION IMPLICATIONS MIXTES (GIM)*

Données : Un contexte formel $R = (G, M, I)$.

Résultat : $\Sigma_{R|\tilde{R}}$.

Le problème GIM est lui aussi équivalent au problème GIP. Pour obtenir l'ensemble des implications de $\Sigma_{R|\tilde{R}}$, il suffit de faire l'apposition des deux contextes formels R et \tilde{R} , et de lui appliquer les algorithmes connus pour le problème GIP.

Problème 12 *GENERATION IMPLICATIONS DUALES (GID)*

Données : Un contexte formel $R = (G, M, I)$ et Σ_R .

Résultat : $\Sigma_{\tilde{R}}$.

Le problème GID prend en entrées le contexte formel $R = (G, M, I)$ ainsi que l'ensemble des implications positives associées afin de calculer l'ensemble des implications négatives $\Sigma_{\tilde{R}}$ associé à R . La taille du résultat peut être exponentielle en la taille de la donnée si nous considérons que les bases d'implications sont des bases génériques. Il ne peut donc pas exister d'algorithme polynomial en la taille de l'entrée. Nous montrons dans l'exemple 21 un contexte illustrant cette observation.

Exemple 21 *Considérons le contexte formel $R = (G, M, I)$ avec $G = \{1, \dots, n-1\}$, $M = \{a_1, \dots, a_n\}$ et $(i, a_j) \in I$ si $j = n - i + 1$. Cela signifie en particulier qu'il n'existe pas d'objet ayant la propriété a_1 . Alors les bases génériques respectives de R et \tilde{R} sont données respectivement par :*

- $\Sigma_R = \{a_1 \rightarrow M \setminus \{a_1\}\} \cup \{a_i a_j \rightarrow a_1 \mid i, j \in \{2, \dots, n\}, i \neq j\}$.
- $\Sigma_{\tilde{R}} = \{\tilde{A} \rightarrow \tilde{a}_1 \mid \emptyset \subset \tilde{A} \subseteq \tilde{M} \setminus \{\tilde{a}_1\}\}$.

Alors $|\Sigma_R| = 1 + (n-1)(n-2)/2$ et $|\Sigma_{\tilde{R}}| = 2^{n-1} - 1$.

Problème 13 *GENERATION CONCEPTS NEGATIFS (GCN)*

Données : Un contexte formel $R = (G, M, I)$ et $\mathcal{B}(R)$ le treillis associé à R .

Résultat : $\mathcal{B}(\tilde{R})$ le treillis associé à \tilde{R} .

Le but du problème GCN est de générer le treillis des concepts associé à la relation complémentaire \tilde{R} étant donnés la relation R et le treillis des concepts $\mathcal{B}(R)$ qui lui est associé.

Pour ce problème, la taille du treillis $\mathcal{B}(\tilde{R})$ peut être exponentielle en la taille de $\mathcal{B}(R)$

(voir l'exemple 22). Cependant les concepts de $\mathcal{B}(\tilde{R})$ peuvent être générés en temps total polynomial à partir de R (c'est à dire en temps polynomial par concept) en utilisant un algorithme de génération de concepts [82].

Exemple 22 *Considérons le contexte formel $R = (G, M, I)$ avec $G = \{1, \dots, n\}$, $M = \{a_1, a_2, \dots, a_n\}$ et $(i, a_j) \in I$ si $j = i$. Le contexte R admet $n + 2$ concepts alors que sa relation complémentaire \tilde{R} en comporte 2^n .*

Problème 14 *GENERATION COMPLETE DES IMPLICATIONS DEPUIS CONCEPTS (GCIC)*

Données : *Un contexte formel $R = (G, M, I)$ et $\mathcal{B}(R)$, $\mathcal{B}(\tilde{R})$ les treillis respectivement associés à R et à \tilde{R} .*

Résultat : $\Sigma_{R|\tilde{R}}$.

Le problème GCIC peut être résolu en assemblant d'une part $\mathcal{B}(R)$ et $\mathcal{B}(\tilde{R})$ en une sous-structure du produit direct de ces deux treillis [98]. La taille du treillis résultant $\mathcal{B}(R|\tilde{R})$ est ainsi inférieure ou égale à la taille du produit cartésien des deux treillis de départ. $\Sigma_{R|\tilde{R}}$ se déduit ainsi en temps polynomial à partir de $\mathcal{B}(R|\tilde{R})$.

Corollaire 3 *Le problème GCIC est polynomial.*

Problème 15 *GENERATION COMPLETE IMPLICATIONS BIS (GCIB)*

Données : Σ_R et $\Sigma_{\tilde{R}}$ les bases d'implications génériques respectives de R et \tilde{R} .

Résultat : $\Sigma_{R|\tilde{R}}$.

Le problème GCIB prend en entrée les bases d'implications positives et négatives d'un contexte R afin d'obtenir une base d'implications mixtes de ce contexte.

Ce problème est à ce jour, à notre connaissance, encore ouvert et sera étudié plus en détails dans le prochain chapitre.

La table 6.3 résume les divers problèmes évoqués en terme de données, de résultat et de complexité.

Problème	Données	Résultat	Classe de Complexité	Taille des données
GIP	R	Σ_R	ouvert	exponentielle
GIN	R	$\Sigma_{\tilde{R}}$	ouvert	exponentielle
GIM	R	$\Sigma_{R \tilde{R}}$	ouvert	exponentielle
GID	R, Σ_R	$\Sigma_{\tilde{R}}$	ouvert	exponentielle
GCN	$R, \mathcal{B}(R)$	$\mathcal{B}(\tilde{R})$	temps total polynomial	exponentielle
GCIC	$R, \mathcal{B}(R), \mathcal{B}(\tilde{R})$	$\Sigma_{R \tilde{R}}$	polynomiale	polynomiale
GCIB	$\Sigma_R, \Sigma_{\tilde{R}}$	$\Sigma_{R \tilde{R}}$	ouvert	ouvert

TAB. 6.3 – Liste des problèmes et leur complexité relative

Chapitre 7

Génération d'implications mixtes

Sommaire

7.1	Propriétés et règles d'inférence sur les implications mixtes	84
7.1.1	Propriétés	84
7.1.2	Règles d'inférence	86
7.1.3	Récapitulatif	89
7.2	Complétude de la génération des implications mixtes	90
7.3	Cas particulier : Contexte réduit	92
7.3.1	Contexte complet	93
7.3.2	Nouvelle règle d'inférence	95
7.3.3	Complétude de la génération d'implications mixtes	96
7.3.4	Exemple complet	100
7.4	Conclusion	102

Dans le chapitre précédent, nous avons décrit différents scénarios possibles pour générer des bases d'implications mixtes ainsi que leur complexité connue à ce jour.

Dans le présent chapitre, nous allons nous intéresser plus particulièrement au problème (GCIB), c'est à dire le problème de génération d'une base d'implications mixtes d'un contexte étant données ses bases d'implications positives et négatives Σ_R et $\Sigma_{\bar{R}}$. Dans le reste de ce chapitre, nous supposons que les bases positives et négatives considérées sont des bases génériques. Ce choix est dû au fait que cette base d'implications est uniquement définie et très utilisée dans différents domaines applicatifs.

7.1 Propriétés et règles d'inférence sur les implications mixtes

Nous exhibons dans cette section quelques propriétés sur les implications mixtes ainsi que des règles d'inférences permettant de déduire un ensemble partiel d'implications mixtes seulement à partir des implications positives ou des implications négatives.

Par souci de clarté, nous présentons les différentes propriétés et les différentes règles d'inférence sous la forme de tableaux où nous leur associons un identifiant :

Identifiant	Propriété
P_1	propriété 4
P_2	propriété 5
P_3	propriété 6
P_4	propriété 7

TAB. 7.1 – Liste des propriétés sur les règles mixtes

Identifiant	Proposition/Corollaire
I_1	proposition 13
I_2	corollaire 4
I_3	proposition 14
I_4	corollaire 5
I_5	proposition 15
I_6	proposition 17

TAB. 7.2 – Liste des règles d'inférence sur les implications mixtes

7.1.1 Propriétés

Nous allons voir ici différentes propriétés concernant les implications mixtes à partir des implications positives et négatives.

Propriété 4 Soient $R = (G, M, I)$ un contexte formel, $A \rightarrow B \in \Sigma_R$ et $\tilde{C} \rightarrow \tilde{D} \in \Sigma_{\tilde{R}}$ si et seulement si $\Sigma_{R|\tilde{R}} \vdash A \rightarrow B$ et $\Sigma_{R|\tilde{R}} \vdash \tilde{C} \rightarrow \tilde{D}$.

Nous avons vu dans l'introduction de cette partie que les implications positives et négatives sont des implications mixtes particulières.

Si nous considérons un contexte R , les implications positives associées à R et les implications négatives associées à \tilde{R} sont aussi des implications pour le contexte $R|\tilde{R}$ puisque R et \tilde{R} sont des sous contextes de $R|\tilde{R}$ partageant le même ensemble d'objet. La propriété 4 en découle. Elles représentent donc des implications mixtes pour le contexte R .

Une précision est à apporter sur les implications de support nul. Puisque les implications mixtes sont définies sur l'ensemble des attributs $M \cup \tilde{M}$, les implications positives et négatives restent bien valides mais portent sur ce nouvel ensemble d'attributs. Ainsi une implication de R du type $X \rightarrow M$ deviendra $X \rightarrow M\tilde{M}$ pour le contexte $R|\tilde{R}$.

Propriété 5 *Pour tout $x \in M$, on a $\Sigma_{R|\tilde{R}} \vdash x\tilde{x} \rightarrow M\tilde{M}$ [0].*

Les implications mixtes d'un contexte R correspondent aux implications du contexte résultant de l'apposition du contexte R et de son complémentaire \tilde{R} . Ainsi, tout objet du contexte $R|\tilde{R}$ ne peut contenir à la fois un attribut et son attribut complémentaire. En d'autres termes, un attribut ne peut pas être présent et absent en même temps dans un objet. La propriété 5 découle ainsi de cette constatation.

Propriété 6 *Soit $R = (G, M, I)$ un contexte formel. Si un objet dans R est composé de tous les attributs de M , alors $\forall A \subseteq M$ et $\forall \tilde{x} \in \tilde{M}$, $\Sigma_{R|\tilde{R}} \not\vdash A \rightarrow \tilde{x}$.*

Preuve. Soit $A \subseteq M$ et g un objet de G tel qu'il contienne tous les attributs de M . Il est clair que g ne contient aucun attribut de \tilde{M} et par conséquent $A^{\Sigma_{R|\tilde{R}}} \subseteq M$. ■

Exemple 23 *Nous modifions le contexte formel de l'exemple 6.1 en ajoutant l'objet 5 qui contient tous les attributs de M .*

$R \tilde{R}$	a	b	c	d	\tilde{a}	\tilde{b}	\tilde{c}	\tilde{d}
1	1	0	0	1	0	1	1	0
2	0	1	0	1	1	0	1	0
3	1	1	0	0	0	0	1	1
4	0	1	1	1	1	0	0	0
5	1	1	1	1	0	0	0	0

Nous pouvons ici voir que tout sous-ensemble $A \subseteq M$ ne peut pas impliquer un attribut de \tilde{M} . $\Sigma_{R|\tilde{R}} = \{c \rightarrow bd, ac \rightarrow bd, abd \rightarrow c, \tilde{a} \rightarrow bd, \tilde{b} \rightarrow a\tilde{c}d, a\tilde{c}d \rightarrow \tilde{b}, \tilde{a}\tilde{c} \rightarrow bd, b\tilde{c}d \rightarrow \tilde{a}, \tilde{d} \rightarrow ab\tilde{c}, ab\tilde{c} \rightarrow \tilde{d}, \tilde{a}c \rightarrow bd, c\tilde{d} \rightarrow a\tilde{a}b\tilde{b}\tilde{c}d[0], \tilde{a}\tilde{d} \rightarrow ab\tilde{b}\tilde{c}\tilde{d}[0], \tilde{a}\tilde{b} \rightarrow abc\tilde{c}d\tilde{d}[0], \tilde{b}c \rightarrow a\tilde{a}b\tilde{c}d\tilde{d}[0], \tilde{b}\tilde{d} \rightarrow a\tilde{a}b\tilde{c}\tilde{d}[0], ab\tilde{c}d \rightarrow \tilde{a}\tilde{b}\tilde{c}\tilde{d}[0], a\tilde{a} \rightarrow b\tilde{b}\tilde{c}\tilde{d}\tilde{d}[0], b\tilde{b} \rightarrow a\tilde{a}\tilde{c}\tilde{d}\tilde{d}[0], c\tilde{c} \rightarrow a\tilde{a}b\tilde{b}\tilde{d}\tilde{d}[0], d\tilde{d} \rightarrow a\tilde{a}b\tilde{b}\tilde{c}\tilde{c}[0]\}$

La propriété 6 montre qu'un objet composé de tous les attributs du contexte R peut empêcher la présence d'un type d'implications mixtes, en l'occurrence des implications dont la prémisse est constituée d'attribut de M et dont la conclusion comporte des attributs de \tilde{M} . La présence d'un tel objet dans R se traduit sur sa base générique associée Σ_R par le fait qu'elle ne contient aucune implication dont le support est nul.

Nous pouvons alors appliquer le même raisonnement sur la relation complémentaire \tilde{R} . La propriété 7 en découle.

Propriété 7 *Soit $R = (G, M, I)$ un contexte formel. Si un objet dans R n'est composé d'aucun attribut de M , alors $\forall \tilde{A} \subseteq \tilde{M}$ et $\forall x \in M$, $\Sigma_{R|\tilde{R}} \not\vdash \tilde{A} \rightarrow x$.*

Preuve. Si un objet dans R n'est composé d'aucun attribut de M , alors il existe un objet dans \tilde{R} composé de tous les attributs de \tilde{M} . Ainsi le raisonnement de la preuve est le même que pour la propriété 6. ■

7.1.2 Règles d'inférence

Nous allons voir dans cette section que nous pouvons déduire des implications mixtes à partir de l'ensemble des implications positives et négatives. Ainsi, un sous-ensemble des implications mixtes peut être généré de manière efficace.

Pour illustrer les différentes propositions et résultats que nous avons obtenus sur ce problème, nous considérons l'exemple présenté dans la figure 7.1. Dans cet exemple, nous avons représenté un contexte formel R sous sa forme matricielle ainsi que son contexte complémentaire \tilde{R} et l'apposition de ces deux contextes. Nous donnons aussi les bases génériques associées à chacun de ces contextes. La figure 7.2 représente les systèmes de fermeture associés à ces différents contextes.

Proposition 13 *Soit $R = (G, M, I)$ un contexte formel et $Ax \rightarrow y \in \Sigma_R$. Alors $\Sigma_{R|\tilde{R}} \vdash A\tilde{y} \rightarrow \tilde{x}$.*

Preuve. Nous considérons ici deux cas distincts :

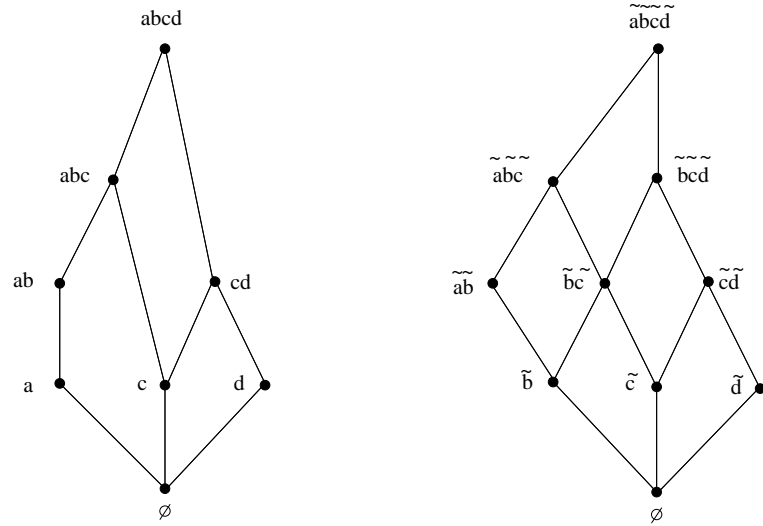
- $\exists F \in \mathcal{F}_{R|\tilde{R}}$ tel que $A\tilde{y} \subseteq F$. Puisque $Ax \rightarrow y$ (par la propriété 4), $x \notin F$ sinon y appartiendrait à F . Ainsi nous avons $\Sigma_{R|\tilde{R}} \vdash A\tilde{y} \rightarrow \tilde{x}$.
- $\nexists F \in \mathcal{F}_{R|\tilde{R}}$ tel que $A\tilde{y} \subseteq F$. Alors $\Sigma_{R|\tilde{R}} \vdash A\tilde{y} \rightarrow M\tilde{M} [0]$ et par conséquent $\Sigma_{R|\tilde{R}} \vdash A\tilde{y} \rightarrow \tilde{x}$.

R	a	b	c	d
1	1	1	1	0
2	1	1	0	0
3	1	0	0	0
4	0	0	0	1
5	0	0	1	1

\tilde{R}	\tilde{a}	\tilde{b}	\tilde{c}	\tilde{d}
1	0	0	0	1
2	0	0	1	1
3	0	1	1	1
4	1	1	1	0
5	1	1	0	0

$$\Sigma_R = \{b \rightarrow a, ac \rightarrow b, bc \rightarrow a, ad \rightarrow bc[0], bd \rightarrow ac[0]\}$$

$$\Sigma_{\tilde{R}} = \{\tilde{a} \rightarrow \tilde{b}, \tilde{b}\tilde{d} \rightarrow \tilde{c}, \tilde{a}\tilde{c} \rightarrow \tilde{b}, \tilde{a}\tilde{d} \rightarrow \tilde{b}\tilde{c}[0]\}$$

FIG. 7.1 – Un contexte formel R et son complémentaire \tilde{R} FIG. 7.2 – Systèmes de fermeture associés à R et \tilde{R} de la figure 7.1

■

La proposition 13 permet ainsi de générer des implications mixtes à partir d'implications positives. Nous pouvons appliquer le même raisonnement à partir des implications négatives puisqu'elles sont obtenues à partir du contexte complémentaire \tilde{R} . Nous déduisons par ce fait le corollaire suivant :

Corollaire 4 *Soit $R = (G, M, I)$ un contexte formel et $\tilde{A}\tilde{x} \rightarrow \tilde{y} \in \Sigma_{\tilde{R}}$. Alors $\Sigma_{R|\tilde{R}} \vdash \tilde{A}y \rightarrow x$.*

Preuve. La preuve de ce corollaire est similaire à celle de la proposition 13 en considérant le système de fermeture $\mathcal{F}_{\tilde{R}}$ associé au contexte \tilde{R} . ■

Exemple 24 *Si nous considérons les bases d'implications positives et négatives de la figure 7.1, nous pouvons ainsi inférer les implications mixtes suivantes qui sont induites par $\Sigma_{R|\tilde{R}}$.*

Implications de Σ_R		Implications induites
$b \rightarrow a$	$\Rightarrow I_1 \Rightarrow$	$\tilde{a} \rightarrow \tilde{b}$
$ac \rightarrow b$	$\Rightarrow I_1 \Rightarrow$	$a\tilde{b} \rightarrow \tilde{c}$
		$\tilde{c}\tilde{b} \rightarrow \tilde{a}$
$bc \rightarrow a$	$\Rightarrow I_1 \Rightarrow$	$c\tilde{a} \rightarrow \tilde{b}$
		$b\tilde{a} \rightarrow \tilde{c}$
Implications de $\Sigma_{\tilde{R}}$		Implications induites
$\tilde{a} \rightarrow \tilde{b}$	$\Rightarrow I_2 \Rightarrow$	$b \rightarrow a$
$\tilde{b}\tilde{d} \rightarrow \tilde{c}$	$\Rightarrow I_2 \Rightarrow$	$\tilde{d}c \rightarrow b$
		$\tilde{b}c \rightarrow d$
$\tilde{a}\tilde{c} \rightarrow \tilde{b}$	$\Rightarrow I_2 \Rightarrow$	$\tilde{c}b \rightarrow a$
		$\tilde{a}b \rightarrow c$

De la même manière que la proposition 13 et le corollaire 4, nous pouvons déduire d'autres implications mixtes à partir des implications positives et négatives ayant un support nul.

Proposition 14 *Soit $R = (G, M, I)$ un contexte formel et $Ax \rightarrow M [0] \in \Sigma_R$. Alors $\Sigma_{R|\tilde{R}} \vdash A \rightarrow \tilde{x}$. De manière duale, si $A \rightarrow \tilde{x} \in \Sigma_{R|\tilde{R}}$ alors $\Sigma_R \vdash Ax \rightarrow M [0]$.*

Preuve. Soit $Ax \rightarrow M [0]$, alors $(Ax)^{\Sigma_{R|\tilde{R}}} = M\tilde{M}$. Alors pour tout $F \in \mathcal{F}_{R|\tilde{R}}$, $F \neq M\tilde{M}$, tel que $A \subseteq F$, $x \notin F$, et par conséquent $\tilde{x} \in F$.

Inversement, si $\Sigma_{R|\tilde{R}} \vdash A \rightarrow \tilde{x}$. Il n'existe pas de $F \in \mathcal{F}_R$ tel que $Ax \subseteq F$ hormis M . ■

Corollaire 5 Soit $R = (G, M, I)$ un contexte formel et $\tilde{A}\tilde{x} \rightarrow \tilde{M} [0] \in \Sigma_{\tilde{R}}$. Alors $\Sigma_{R|\tilde{R}} \vdash \tilde{A} \rightarrow x$. De manière duale, si $\tilde{A} \rightarrow x \in \Sigma_{R|\tilde{R}}$ alors $\Sigma_{\tilde{R}} \vdash \tilde{A}\tilde{x} \rightarrow \tilde{M} [0]$

Preuve. Le raisonnement de la preuve est le même que la preuve de la proposition 14. ■

Exemple 25 Si nous considérons les bases d'implications positives et négatives de la figure 7.1. Nous pouvons ainsi inférer les implications mixtes suivantes qui sont induites par $\Sigma_{R|\tilde{R}}$.

$$\begin{array}{ccc}
 \text{Implications de } \Sigma_R & & \text{Implications induites} \\
 ad \rightarrow bc [0] & \Rightarrow I_3 \Rightarrow & a \rightarrow \tilde{d} \\
 & & d \rightarrow \tilde{a} \\
 bd \rightarrow ac [0] & \Rightarrow I_3 \Rightarrow & b \rightarrow \tilde{d} \\
 & & d \rightarrow \tilde{b} \\
 \\
 \text{Implications de } \Sigma_{\tilde{R}} & & \text{Implications induites} \\
 \tilde{a}\tilde{d} \rightarrow \tilde{b}\tilde{c} & \Rightarrow I_4 \Rightarrow & \tilde{a} \rightarrow d \\
 & & \tilde{d} \rightarrow a
 \end{array}$$

La proposition 14 et le corollaire 5 peuvent ainsi être généralisés et exprimés sous forme d'équivalence d'implications comme suit :

Proposition 15 L'implication $A \rightarrow \tilde{x} \in \Sigma_{R|\tilde{R}}$ pour $A \subseteq MM$ est équivalente à $Ax \rightarrow MM [0]$.

Preuve. Cette proposition peut ainsi être prouvée en conduisant le même raisonnement que celui utilisé pour la Proposition 13. ■

7.1.3 Récapitulatif

Nous pouvons voir dans la table 7.3 un récapitulatif des différentes propositions évoquées dans ce chapitre.

Ceci nous permet ainsi de générer un ensemble d'implications mixtes d'un contexte donné R à partir de ses bases d'implications génériques positives et négatives Σ_R et $\Sigma_{\tilde{R}}$. Malheureusement, ces ensembles d'implications ne définissent pas dans tous les cas une base d'implications pour le contexte $R|\tilde{R}$.

Implications considérées		Implications déduites
$Ax \rightarrow y \in \Sigma_R$	$\Rightarrow I_1 \Rightarrow$	$A\tilde{y} \rightarrow \tilde{x}$
$\tilde{A}\tilde{x} \rightarrow \tilde{y} \in \Sigma_{\tilde{R}}$	$\Rightarrow I_2 \Rightarrow$	$\tilde{A}y \rightarrow x$
$Ax \rightarrow M[0] \in \Sigma_R$	$\Rightarrow I_3 \Rightarrow$	$A \rightarrow \tilde{x}$
$\tilde{A}\tilde{x} \rightarrow \tilde{M}[0] \in \Sigma_{\tilde{R}}$	$\Rightarrow I_4 \Rightarrow$	$\tilde{A} \rightarrow x$

TAB. 7.3 – Récapitulatif des différentes propositions sur les implications mixtes

7.2 Complétude de la génération des implications mixtes

Nous allons voir dans cette section qu'il est impossible de calculer l'ensemble complet des implications mixtes uniquement à partir des bases d'implications positives et négatives d'un contexte R .

Ceci vient du fait que l'on peut associer à un couple de bases d'implications positives et négatives Σ_R et $\Sigma_{\tilde{R}}$ différents contextes formels qui respectent ces bases d'implications. L'exemple qui suit présente le cas où nous considérons deux contextes R_1 et R_2 tel que leurs bases d'implications positives et négatives respectives sont équivalentes et tel que leurs bases d'implications mixtes sont différentes.

Exemple 26 *Considérons les contextes formels R_1 et R_2 de la figure 7.3. Nous pouvons voir que R_1 est un contexte réduit de R_2 puisque l'objet 6 dans R_2 peut être obtenu par l'intersection de l'objet 3 et l'objet 4. De plus puisque l'objet 6 est l'union de l'objet 1 et de l'objet 2, l'objet 6 dans \tilde{R}_2 devient ainsi l'intersection de l'objet 1 et de l'objet 2. \tilde{R}_1 est ainsi un contexte réduit de \tilde{R}_2 . Les bases d'implications positives et négatives de R_1 et R_2 sont donc équivalentes entre elles.*

Or si nous considérons les bases d'implications génériques associées à $R_1|\tilde{R}_1$ et $R_2|\tilde{R}_2$, elles sont différentes (figure 7.4). Ainsi, si nous considérons l'ensemble $ab\tilde{c}\tilde{d}$, nous avons $(ab\tilde{c}\tilde{d})^{\Sigma_{R_1}|\tilde{R}_1} = ab\tilde{c}\tilde{d}\tilde{e}$ et $(ab\tilde{c}\tilde{d})^{\Sigma_{R_2}|\tilde{R}_2} = ab\tilde{c}\tilde{d}$.

La proposition 16 découle ainsi de l'exemple 26.

Proposition 16 *Soit Σ_R et $\Sigma_{\tilde{R}}$ des bases d'implications pour un contexte formel R et son complémentaire \tilde{R} . Il n'est pas possible de générer une base d'implications complète pour le contexte $R|\tilde{R}$.*

Nous ne pouvons donc pas générer une base d'implications mixtes complète pour un contexte R si les seules données dont nous disposons sont ses bases d'implications gén-

R_1	a	b	c	d	e
1	1	1	0	0	0
2	1	0	0	0	1
3	1	1	1	0	1
4	1	1	0	1	1
5	0	1	0	0	1

 R_1

R_2	a	b	c	d	e
1	1	1	0	0	0
2	1	0	0	0	1
3	1	1	1	0	1
4	1	1	0	1	1
5	0	1	0	0	1
6	1	1	0	0	1

 R_2 FIG. 7.3 – R_1 est un contexte réduit de R_2

$\tilde{e} \rightarrow ab\tilde{c}\tilde{d}$	$b\tilde{b} \rightarrow a\tilde{a}c\tilde{c}d\tilde{d}e\tilde{e}[0]$
$ab\tilde{c}\tilde{d} \rightarrow \tilde{e}$	$c\tilde{b} \rightarrow a\tilde{a}b\tilde{c}d\tilde{d}e\tilde{e}[0]$
$\tilde{b} \rightarrow ae\tilde{c}\tilde{d}$	$d\tilde{b} \rightarrow a\tilde{a}bc\tilde{c}d\tilde{d}e\tilde{e}[0]$
$ae\tilde{c}\tilde{d} \rightarrow \tilde{b}$	$e\tilde{e} \rightarrow a\tilde{a}b\tilde{b}c\tilde{c}d\tilde{d}[0]$
$c \rightarrow abe\tilde{d}$	$a\tilde{a} \rightarrow b\tilde{b}c\tilde{c}d\tilde{d}e\tilde{e}[0]$
$abe\tilde{d} \rightarrow c$	$\tilde{a}\tilde{b} \rightarrow abc\tilde{c}d\tilde{d}e\tilde{e}[0]$
$d \rightarrow abe\tilde{c}$	$c\tilde{a} \rightarrow ab\tilde{b}c\tilde{c}d\tilde{d}e\tilde{e}[0]$
$abe\tilde{c} \rightarrow d$	$d\tilde{a} \rightarrow ab\tilde{b}c\tilde{c}d\tilde{d}e\tilde{e}[0]$
$\tilde{a} \rightarrow be\tilde{c}\tilde{d}$	$\tilde{a}\tilde{e} \rightarrow ab\tilde{b}c\tilde{c}d\tilde{d}e[0]$
$be\tilde{c}\tilde{d} \rightarrow \tilde{a}$	$c\tilde{e} \rightarrow a\tilde{a}b\tilde{b}c\tilde{c}d\tilde{d}e[0]$
$c\tilde{c} \rightarrow a\tilde{a}b\tilde{b}d\tilde{d}e\tilde{e}[0]$	$\tilde{b}\tilde{e} \rightarrow a\tilde{a}bc\tilde{c}d\tilde{d}e[0]$
$c\tilde{d} \rightarrow a\tilde{a}b\tilde{b}c\tilde{d}e\tilde{e}[0]$	$d\tilde{e} \rightarrow a\tilde{a}b\tilde{b}c\tilde{c}d\tilde{e}[0]$
$d\tilde{d} \rightarrow a\tilde{a}b\tilde{b}c\tilde{c}e\tilde{e}[0]$	$abe\tilde{c}\tilde{d} \rightarrow \tilde{a}\tilde{b}cd\tilde{e}[0]$

 $\Sigma_{R_1|\tilde{R}_1}$

$\tilde{e} \rightarrow ab\tilde{c}\tilde{d}$	$d\tilde{e} \rightarrow a\tilde{a}b\tilde{b}c\tilde{c}d\tilde{e}[0]$
$\tilde{b} \rightarrow ae\tilde{c}\tilde{d}$	$e\tilde{e} \rightarrow a\tilde{a}b\tilde{b}c\tilde{c}d\tilde{d}[0]$
$c \rightarrow abe\tilde{d}$	$c\tilde{a} \rightarrow ab\tilde{b}c\tilde{c}d\tilde{d}e\tilde{e}[0]$
$d \rightarrow abe\tilde{c}$	$d\tilde{a} \rightarrow ab\tilde{b}c\tilde{c}d\tilde{d}e\tilde{e}[0]$
$\tilde{a} \rightarrow be\tilde{c}\tilde{d}$	$\tilde{a}\tilde{e} \rightarrow ab\tilde{b}c\tilde{c}d\tilde{d}e[0]$
$a\tilde{a} \rightarrow b\tilde{b}c\tilde{c}d\tilde{d}e\tilde{e}[0]$	$\tilde{a}\tilde{b} \rightarrow abc\tilde{c}d\tilde{d}e\tilde{e}[0]$
$d\tilde{d} \rightarrow a\tilde{a}b\tilde{b}c\tilde{c}e\tilde{e}[0]$	$b\tilde{b} \rightarrow a\tilde{a}c\tilde{c}d\tilde{d}e\tilde{e}[0]$
$c\tilde{c} \rightarrow a\tilde{a}b\tilde{b}d\tilde{d}e\tilde{e}[0]$	$\tilde{b}\tilde{e} \rightarrow a\tilde{a}bc\tilde{c}d\tilde{d}e[0]$
$cd \rightarrow a\tilde{a}b\tilde{b}c\tilde{d}e\tilde{e}[0]$	$c\tilde{b} \rightarrow a\tilde{a}b\tilde{c}d\tilde{d}e\tilde{e}[0]$
$c\tilde{e} \rightarrow a\tilde{a}b\tilde{b}c\tilde{d}d\tilde{e}[0]$	$d\tilde{b} \rightarrow a\tilde{a}bc\tilde{c}d\tilde{d}e\tilde{e}[0]$

 $\Sigma_{R_2|\tilde{R}_2}$ FIG. 7.4 – Bases d'implications génériques pour $R_1|\tilde{R}_1$ et $R_2|\tilde{R}_2$ de la figure 7.3

riques positives et négatives. Ceci est en partie du au fait qu'une base d'implications peut représenter **plusieurs contextes**.

7.3 Cas particulier : Contexte réduit

Nous nous concentrons dans cette section à une restriction des contextes généraux : les contextes réduits. L'intérêt principal de ces contextes est leur unicité. Une base d'implications ne peut ainsi représenter qu'un seul et unique contexte réduit. Malheureusement, les propriétés et les règles d'inférence que nous avons énoncées dans la section 7.1.2 ne permettent pas une génération complète d'une base d'implications mixtes à partir des bases d'implications positives et négatives d'un contexte R réduit. L'exemple 27 représente ainsi un contexte R réduit ainsi que ses bases d'implications génériques positives et négatives associées. Nous pouvons voir que l'implication mixte $ab\tilde{c}\tilde{d} \rightarrow cde\tilde{a}\tilde{b}\tilde{e}[0]$ est valide pour le contexte $R|\tilde{R}$. Or nous ne pouvons pas l'obtenir à partir des bases d'implications positives et négatives de R en utilisant les propriétés et règles d'inférence évoquées précédemment.

Exemple 27 Soit R un contexte formel décrit ci-dessous avec ses bases d'implications génériques positives et négatives associées :

R	a	b	c	d	e
1	1	1	0	1	0
2	1	1	1	0	0
3	0	1	0	0	1
4	1	0	0	0	0

$$\begin{aligned}\Sigma_R &= \{e \rightarrow b, d \rightarrow ab, c \rightarrow ab, cd \rightarrow abe[0], ae \rightarrow bcd[0], ce \rightarrow abd[0], de \rightarrow abc[0]\} \\ \Sigma_{\tilde{R}} &= \{\tilde{a} \rightarrow \tilde{c}\tilde{d}, \tilde{b} \rightarrow \tilde{c}\tilde{d}\tilde{e}, \tilde{c}\tilde{d}\tilde{e} \rightarrow \tilde{b}, \tilde{a}\tilde{b} \rightarrow \tilde{c}\tilde{d}\tilde{e}[0], \tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]\}\end{aligned}$$

Nous introduisons ainsi dans cette section une nouvelle règle d'inférence spécifique aux contextes réduits. Cette dernière nous permet de montrer, avec l'aide des précédentes règles d'inférence, que nous pouvons générer une base d'implications mixtes valide et complète d'un contexte réduit R à partir de ses bases d'implications positives et négatives associées. Pour arriver à ce résultat, nous traiterons un cas particulier que nous appellerons “contexte complet” indépendamment des autres cas. Un “contexte complet” se traduit par le fait que l'une de ces deux bases d'implications (positives ou négatives) est vide, c'est à dire ne contient aucune implication. Si nous l'exprimons sur le système de fermeture associé, cela signifie que toutes les parties de G sont des ensembles fermés.

7.3.1 Contexte complet

Il existe des contextes formels tel que leur base d'implications positives associée (ou bien leur base d'implications négatives) est vide. Il existe ainsi un unique contexte réduit qui représente cette base d'implications positive vide (le contexte complémentaire pour la base d'implications négatives vide). Les objets de ce contexte R contiennent tous

R	a	b	c	d
1	0	1	1	1
2	1	0	1	1
3	1	1	0	1
4	1	1	1	0

$$\Sigma_R = \{\}$$

$$\Sigma_{\tilde{R}} = \{\tilde{a}\tilde{b} \rightarrow \tilde{c}\tilde{d}[0], \tilde{a}\tilde{c} \rightarrow \tilde{b}\tilde{d}[0], \tilde{a}\tilde{d} \rightarrow \tilde{b}\tilde{c}[0], \tilde{b}\tilde{c} \rightarrow \tilde{a}\tilde{d}[0], \tilde{b}\tilde{d} \rightarrow \tilde{a}\tilde{c}[0], \tilde{c}\tilde{d} \rightarrow \tilde{a}\tilde{b}[0]\}$$

FIG. 7.5 – R un contexte complet à 4 attributs et ses bases d'implications génériques positives et négatives associées

les attributs de R hormis un. La taille de l'ensemble des objets est égale à la taille de l'ensemble des attributs puisque tous les objets sont caractérisés par l'absence de chacun des attributs. La figure 7.5 représente ainsi un contexte complet à quatre attributs avec ses bases d'implications positives et négatives associées.

Si nous considérons un contexte complet $R=(G, M, I)$ de taille k où $M = (a_1, \dots, a_k)$, les bases d'implications positives et négatives associées sont construites de la manière suivante :

$$- \Sigma_R = \{\}$$

$$- \Sigma_{\tilde{R}} = \{\tilde{a}_i\tilde{a}_j \rightarrow \tilde{M} \setminus \{\tilde{a}_i\tilde{a}_j\}[0], \forall i \in \{1..k\}, \forall j \in \{1..k\}, i \neq j\}$$

Par la construction du contexte complet R , la base d'implications mixtes générique est construite de la manière suivante lorsque l'ensemble des attributs est de taille k :

$$\begin{aligned} \Sigma_{R|\tilde{R}} = & \quad \{\tilde{a}_i\tilde{a}_j \rightarrow M\tilde{M}[0], \forall i \in \{1..k\}, \forall j \in \{1..k\}, i \neq j\} \\ & \cup \quad \{M \setminus a_i \rightarrow \tilde{a}_i, \forall i \in \{1..k\}\} \\ & \cup \quad \{\tilde{a}_i \rightarrow a_j, \forall i \in \{1..k\}, \forall j \in \{1..k\}, i \neq j\} \\ & \cup \quad a_1 \dots a_k \rightarrow \tilde{a}_1 \dots \tilde{a}_k[0] \\ & \cup \quad \{a_i\tilde{a}_i \rightarrow M\tilde{M}[0], \forall i \in \{1..k\}\} \end{aligned}$$

Or il n'est pas possible, à partir des bases d'implications positives et négatives et des différentes propriétés et règles d'inférence que nous avons proposées, d'obtenir une base d'implications mixtes complète pour un contexte complet.

Ceci vient du fait que la base d'implications positives d'un contexte complet est vide. Nous ne pouvons donc pas générer les règles mixtes du type " $M \setminus a_i \rightarrow \tilde{a}_i$ " avec les propriétés et règles d'inférence que nous avons.

Nous pouvons caractériser la base d'implications mixtes d'un contexte complet étant donné la taille de l'ensemble des attributs du contexte considéré. Ainsi, si nous avons une base d'implications positives vide et le nombre d'attributs d'un contexte réduit, nous pouvons déduire sa base d'implications mixtes de part sa construction.

D'autre part, nous avons remarqué qu'un contexte complet de taille k possède en réalité une implication positive induite :

$$a_1 \dots a_k \rightarrow \{\} [0]$$

Cette implication n'est jamais prise en compte dans les bases d'implications positives classiques car, par elle seule, elle est obsolète par rapport au contexte complet.

En revanche, nous nous sommes rendu compte qu'elle permet de déterminer une base d'implications mixtes pour les contextes complets en utilisant les propriétés et règles d'inférence de la section précédente. Si nous ajoutons cette nouvelle implication à la base d'implications positives du contexte complet, nous obtenons les bases d'implications positives et négatives suivantes :

- $\Sigma_R = \{a_1 \dots a_k \rightarrow \{\} [0]\}$
- $\Sigma_{\tilde{R}} = \{\tilde{a}_i \tilde{a}_j \rightarrow \tilde{M} \setminus \{\tilde{a}_i \tilde{a}_j\} [0], \forall i \in \{1..k\}, \forall j \in \{1..k\}, i \neq j\}$

En appliquant les différentes règles d'inférence sur ces deux bases d'implications nous obtenons les implications mixtes suivantes :

Implications de Σ_R		Implications induites	
$a_1 \dots a_k \rightarrow \{\} [0]$	$\Rightarrow I_3 \Rightarrow$	$M \setminus a_i \rightarrow \tilde{a}_i, \forall i \in \{1..k\}$	
Implications de $\Sigma_{\tilde{R}}$		Implications induites	
$\tilde{a}_i \tilde{a}_j \rightarrow \tilde{M} [0]$	$\Rightarrow I_4 \Rightarrow$	$\tilde{a}_i \rightarrow a_j$	

En utilisant la propriété 4 nous retrouvons toutes les implications positives et négatives plus les implications de la forme $a_i \tilde{a}_i \rightarrow M \tilde{M} [0]$ grâce à la propriété 5.

Nous obtenons ainsi toutes les implications mixtes d'un contexte complet de taille k .

7.3.2 Nouvelle règle d'inférence

Nous avons trouvé un exemple (exemple 27) tel que nous ne pouvons pas inférer toutes les implications pour un contexte $R=(G, M, I)$ réduit en utilisant les règles d'inférence et les propriétés étudiées précédemment.

Les règles d'inférence et les propriétés que nous avons actuellement ne nous suffisent donc pas à la génération d'une base d'implications mixtes d'un contexte réduit R étant donné ses bases d'implications génériques positives et négatives.

Nous avons aussi obtenu une nouvelle règle d'inférence s'appliquant sur les contextes réduits :

Proposition 17 *Soit $R=(G, M, I)$ un contexte réduit. Alors $\Sigma_{R|\tilde{R}} \vdash A \rightarrow x$, $A \in M\tilde{M}$ avec $A = A^+ \cup A^-$ tel que $A^+ \subseteq M$ et $A^- \subseteq \tilde{M}$, $x \in M$ si et seulement $\forall y \in M \setminus A^+$, $\Sigma_{R|\tilde{R}} \vdash Ay \rightarrow x$ et A^+ n'est pas un inf-irréductible de \mathcal{F}_R .*

Preuve. Nous allons donc prouver les deux sens.

- \Rightarrow : Supposons que $\Sigma_{R|\tilde{R}} \vdash A \rightarrow x$. Par la règle d'augmentation d'Armstrong nous avons $\forall y \in M \setminus A$, $\Sigma_{R|\tilde{R}} \vdash Ay \rightarrow x$.
- \Leftarrow : Supposons que $\forall y \in M \setminus A^+$, $\Sigma_{R|\tilde{R}} \vdash Ay \rightarrow x$ et A^+ n'est pas un élément inf-irréductible de \mathcal{F}_R . Deux cas sont alors possibles :
 - Il n'existe pas d'objet $g \in G$ tel que les attributs de A appartiennent à g . Par définition, nous avons $\Sigma_{R|\tilde{R}} \vdash A \rightarrow M\tilde{M}[0]$ et donc $\Sigma_{R|\tilde{R}} \vdash A \rightarrow x \in G$.
 - Il existe au moins un objet g tel que les attributs de A appartiennent à g . Or, pour tout objet $g \in G$ tel que les attributs de A appartiennent à g , il existe $z \in M \setminus A^+$ tel que z appartient à g puisque A^+ n'est pas un inf-irréductible de \mathcal{F}_R . D'après l'hypothèse, $\forall y \in M \setminus A^+$, $\Sigma_{R|\tilde{R}} \vdash Ay \rightarrow x$, nous avons donc $\Sigma_{R|\tilde{R}} \vdash A \rightarrow x$.

■

La condition sur le fait qu'un ensemble ne soit pas un élément inf-irréductible d'un système de fermeture peut être testé à partir d'une base d'implications lui étant associée. Cette inférence peut donc être réalisée en ne considérant que les bases d'implications génériques positives et négatives d'un contexte réduit. Cette nouvelle règle d'inférence nous permet ainsi de trouver de nouvelles implications mixtes pour un contexte réduit donné, que nous n'arrivions pas à obtenir avec les règles d'inférences précédentes. Ainsi, pour l'exemple 27, l'application de cette règle permet de trouver l'implication $abc\tilde{d} \rightarrow M\tilde{M}[0]$. Nous avons ici $A^+ = ab$ et $A^- = \tilde{c}\tilde{d}$ et A^+ n'est pas un inf-irréductible de \mathcal{F}_R .

Nous regardons alors si, pour tout y appartenant à $M \setminus A^+$, $\Sigma_{R|\tilde{R}} \vdash aby\tilde{c}\tilde{d} \rightarrow M\tilde{M}[0]$. Nous obtenons les cas suivants :

- $\Sigma_{R|\tilde{R}} \vdash abc\tilde{c}\tilde{d} \rightarrow M\tilde{M}[0]$ puisque $\Sigma_{R|\tilde{R}} \vdash c\tilde{c} \rightarrow M\tilde{M}[0]$ d'après P_2 .
- $\Sigma_{R|\tilde{R}} \vdash abd\tilde{c}\tilde{d} \rightarrow M\tilde{M}[0]$ puisque $\Sigma_{R|\tilde{R}} \vdash d\tilde{d} \rightarrow M\tilde{M}[0]$ d'après P_2 .
- $\Sigma_{R|\tilde{R}} \vdash abe\tilde{c}\tilde{d} \rightarrow M\tilde{M}[0]$ puisque $\Sigma_{R|\tilde{R}} \vdash ae \rightarrow M\tilde{M}[0]$ du fait que $ae \rightarrow M[0]$ appartient à Σ_R et de la propriété P_1 .

En appliquant la proposition 17, nous aboutissons au fait que $\Sigma_{R|\tilde{R}} \vdash ab\tilde{c}\tilde{d} \rightarrow M\tilde{M}[0]$.

Nous allons voir dans la section suivante que cette règle permet d'obtenir un ensemble valide et complet d'implications mixtes pour un contexte réduit R étant données ses bases d'implications génériques positives et négatives associées.

7.3.3 Complétude de la génération d'implications mixtes

Dans cette section, nous allons voir que les différentes propriétés et les différentes règles d'inférence permettent, à partir des bases d'implications génériques positives et négatives d'un contexte réduit R , d'obtenir un ensemble d'implications mixtes pour R valide et complet.

Dans la suite de cette section, nous supposons que le contexte pris en considération n'est pas un contexte complet et donc qu'il existe au moins une implication contenue dans Σ_R . Nous enonçons ci-dessous le problème associé :

Problème 16 INFERENCE :

Données : Soient Σ_R et $\Sigma_{\tilde{R}}$ les bases d'implications génériques positives et négatives d'un contexte réduit R .

Résultats : $\Sigma_{R|\tilde{R}}$

Nous allons voir, avec le théorème 20, que nous pouvons répondre au problème INFERENCE grâce aux différentes règles d'inférence que nous avons proposées. Afin de démontrer la validité et la complétude de la base d'implications inférée nous exhibons la propriété 8 ainsi que les lemmes 1 et 2. Nous notons ainsi par Σ_{mixte} l'ensemble des implications obtenu par les règles d'inférence et les propriétés définies dans ce chapitre. Un rappel des différentes propriétés et règles d'inférence est visible dans la table 7.4.

Propriété 8 Soit $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x$ tels que $F = F^+ \cup F^-$ avec $F^+ \subseteq M$ et $F^- \subseteq \tilde{M}$ et $\forall z \in F^+, \tilde{z} \notin F^-$. Alors $\forall F' \subseteq M$ tel que $F^+ \subseteq F'$, $x \notin F'$ et $\forall y \in F' \setminus F^+, y \notin F^-$, on a $F' \notin \mathcal{M}(\mathcal{F}_R)$.

Identifiants	Propriétés et Règles d'inférence
P_1	$\Sigma_R \vdash A \rightarrow x \Leftrightarrow \Sigma_{mixte} \vdash A \rightarrow x$
P_1	$\Sigma_{\tilde{R}} \vdash \tilde{A} \rightarrow \tilde{x} \Leftrightarrow \Sigma_{mixte} \vdash \tilde{A} \rightarrow \tilde{x}$
P_1	$\Sigma_R \vdash A \rightarrow M[0] \Leftrightarrow \Sigma_{mixte} \vdash A \rightarrow M\tilde{M}[0]$
P_1	$\Sigma_{\tilde{R}} \vdash \tilde{A} \rightarrow \tilde{M}[0] \Leftrightarrow \Sigma_{mixte} \vdash \tilde{A} \rightarrow M\tilde{M}[0]$
P_2	$\Sigma_{mixte} \vdash a\tilde{a} \rightarrow M\tilde{M}, \forall a \in M$
I_1	$\Sigma_R \vdash Ax \rightarrow y \Rightarrow \Sigma_{mixte} \vdash A\tilde{y} \rightarrow \tilde{x}$
I_2	$\Sigma_{\tilde{R}} \vdash \tilde{A}\tilde{x} \Rightarrow \tilde{y} \Rightarrow \tilde{A}y \rightarrow x$
I_3	$\Sigma_R \vdash Ax \rightarrow M[0] \Rightarrow \Sigma_{mixte} \vdash A \rightarrow \tilde{x}$
I_4	$\Sigma_{\tilde{R}} \vdash \tilde{A}\tilde{x} \rightarrow \tilde{M}[0] \Rightarrow \Sigma_{mixte} \vdash \tilde{A} \rightarrow x$
I_5	$\Sigma_{mixte} \vdash A \rightarrow \tilde{x}, A \subseteq M\tilde{M} \Leftrightarrow \Sigma_{mixte} \vdash Ax \rightarrow M\tilde{M}[0]$
I_6	$\Sigma_{mixte} \vdash A \rightarrow x, A \subseteq M\tilde{M}$ avec $A = A^+ \cup A^-, A^+ \subseteq M, A^- \subseteq \tilde{M}$ et $x \in M$ \Leftrightarrow $\forall y \in M \setminus A^+, \Sigma_{mixte} \vdash Ay \rightarrow x$ et A^+ n'est pas un inf-irréductible de \mathcal{F}_R

TAB. 7.4 – Récapitulatif des différentes propriétés et règles d'inférence engendrant les implications de Σ_{mixtes}

Preuve. Nous prouvons cette propriété par contradiction.

Supposons que $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x$ tels que $F = F^+ \cup F^-$ avec $F^+ \subseteq M$ et $F^- \subseteq \tilde{M}$ et qu'il existe $F' \in \mathcal{M}(\mathcal{F}_R)$ tel que $F^+ \subseteq F'$ et $x \notin F'$. Alors il existe un objet $g \in G$ du contexte $R|\tilde{R}$ tel que les attributs de $F^+ \cup F^-$ sont contenus dans g avec x n'est pas contenu dans g . Cela contredit le fait que $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x$. ■

Lemme 1 Soient $F \subseteq M\tilde{M}$ et $x \in M$. Alors $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x$ implique $\Sigma_{mixte} \vdash F \rightarrow x$.

Preuve. Soit $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x$. Nous considérons ici que $F = F^+ \cup F^-$ avec $F^+ \subseteq M$ et $F^- \subseteq \tilde{M}$. Si $\exists z \in F^+$ tel que $\tilde{z} \in F^-$, d'après P_2 $\Sigma_{mixte} \vdash z\tilde{z} \rightarrow M\tilde{M}[0]$ et donc $\Sigma_{mixte} \vdash F \rightarrow x$. Dans la suite de cette preuve, nous considérons donc que $\forall z \in F^+, \tilde{z} \notin F^-$. Deux cas sont alors possibles :

- Soit $F = F^+$, alors il existe $A \rightarrow x \in \Sigma_{R|\tilde{R}}$ tel que $A \subseteq F^+$ et tel que $x \notin F^+$. D'après la propriété P_1 nous avons $\Sigma_R \vdash A \rightarrow x$ et donc $\Sigma_{mixte} \vdash A \rightarrow x$ d'où $\Sigma_{mixte} \vdash F \rightarrow x$.

- Soit $F = F^+ \cup F^-$, tels que $F^+ \neq \{\}$, $F^- \neq \{\}$. Deux cas sont alors possibles :

1. Soit F^+ n'est pas un inf-irréductible de \mathcal{F}_R . D'après la règle d'inférence I_6 , $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x$, $F^+ \cup F^-$, $x \in M$ si et seulement si $\forall y \in M \setminus F^+$, $\Sigma_{R|\tilde{R}} \vdash Fy \rightarrow x$ et F^+ n'est pas un inf-irréductible de \mathcal{F}_R . Il faut ainsi montrer que $\Sigma_{mixte} \vdash F \rightarrow x$. Pour cela nous regardons si $\forall y \in M \setminus F^+$, $\Sigma_{mixte} \vdash Fy \rightarrow x$. Deux cas sont alors possibles :

- (a) Soit $\tilde{y} \in F^-$. Alors $\Sigma_{mixte} \vdash Fy \rightarrow x$ puisque d'après P_2 , $\Sigma_{mixte} \vdash y\tilde{y} \rightarrow M\tilde{M}[0]$ et par augmentation $Fy \rightarrow M\tilde{M}$ et donc $\Sigma_{mixte} \vdash Fy \rightarrow x$.

- (b) Soit $\tilde{y} \notin F^-$. Deux cas sont alors possibles :

- i. Soit $F^+ \cup y \in \mathcal{M}(\mathcal{F}_R)$. $F^+ \cup y$ ne peut pas être un inf-irréductible de \mathcal{F}_R si $x \notin F^+ \cup y$ et $\tilde{y} \notin F^-$ d'après la propriété 8. F^+ ne peut être un inf-irréductible que si $x \in F^+$. Dans ce cas, $\Sigma_{mixte} \vdash F \cup y \rightarrow x$ puisque $x \in F \cup y$.
- ii. Soit $F^+ \cup y \notin \mathcal{M}(\mathcal{F}_R)$. Il faut alors vérifier que $\Sigma_{mixte} \vdash F \cup y \rightarrow x$ en recommençant le même raisonnement. A chaque itération, la prémisse de l'implication à tester est augmentée d'un élément de M qui n'est pas déjà présent dans celle-ci. La dernière itération de ce processus correspond donc à tester si $\Sigma_{mixte} \vdash M \cup F^- \rightarrow x$. Or, $\exists z \in M$ tel que $\tilde{z} \in F^-$. D'après la propriété P_2 $\Sigma_{mixte} \vdash z\tilde{z} \rightarrow M\tilde{M}[0]$. Par

augmentation, nous avons donc $\Sigma_{mixte} \vdash M \cup F^- \rightarrow x$ et par la règle I_6 , $\Sigma_{mixte} \vdash F \rightarrow x$.

2. Soit F^+ est un inf-irréductible de \mathcal{F}_{Σ_R} . F^+ ne peut pas être un inf-irréductible si $x \notin F^+$ et $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x$ d'après la propriété 8. F^+ ne peut être un inf-irréductible que si $x \in F^+$. Dans ce cas, $\Sigma_{mixte} \vdash F \rightarrow x$ puisque $x \in F$

■

Le dual du lemme 1 est faux puisque le contexte \tilde{R} n'est pas toujours réduit. Nous exhibons ainsi le Lemme 2.

Lemme 2 Soient $F \subseteq M\tilde{M}$ et $\tilde{x} \in M$. Alors $\Sigma_{R|\tilde{R}} \vdash F \rightarrow \tilde{x}$ implique $\Sigma_{mixte} \vdash F \rightarrow \tilde{x}$.

Preuve. Soit $\Sigma_{R|\tilde{R}} \vdash F \rightarrow \tilde{x}$. Nous considérons ici que $F = F^+ \cup F^-$ avec $F^+ \subseteq M$ et $F^- \subseteq \tilde{M}$. Deux cas sont alors possibles :

1. Soit $F = F^-$, alors il existe $\tilde{A} \rightarrow \tilde{x} \in \Sigma_{R|\tilde{R}}$ tel que $\tilde{A} \subseteq F^-$ et tel que $\tilde{x} \notin F^-$. D'après la propriété P_1 nous avons $\Sigma_{\tilde{R}} \vdash \tilde{A} \rightarrow \tilde{x}$ et donc $\Sigma_{mixte} \vdash \tilde{A} \rightarrow \tilde{x}$. Alors $\Sigma_{mixte} \vdash F \rightarrow \tilde{x}$.
2. Soit $F = F^+ \cup F^-$, tel que $F^+ \neq \{\}$, $F^- \neq \{\}$. Deux cas sont alors possibles :
 - (a) Il existe $y \in F^+$ tel que $\tilde{y} \in F^-$. D'après la propriété P_2 $\Sigma_{mixte} \vdash y\tilde{y} \rightarrow M\tilde{M}[0]$. Par augmentation $\Sigma_{mixte} \vdash F \rightarrow M\tilde{M}[0]$ et donc $\Sigma_{mixte} \vdash F \rightarrow \tilde{x}$.
 - (b) Il n'existe pas $x \in F^+$ tel que $\tilde{x} \in F^-$. D'après la règle I_5 , $\Sigma_{R|\tilde{R}} \vdash F \rightarrow \tilde{x}$ si et seulement si $\Sigma_{R|\tilde{R}} \vdash F \cup y \rightarrow M\tilde{M}[0]$. Alors, il existe $\tilde{z} \in F^-$ avec $z \in M$ tel que $\Sigma_{R|\tilde{R}} \vdash F \cup y \rightarrow z$. D'après le lemme 1, $\Sigma_{mixte} \vdash F \cup y \rightarrow z$. Par conséquence, $\Sigma_{mixte} \vdash F \rightarrow \tilde{x}$.

■

Théorème 20 Soient $F \subseteq M\tilde{M}$ et $x \in M\tilde{M}$. Alors $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x$ si et seulement si $\Sigma_{mixte} \vdash F \rightarrow x$.

Preuve. Il nous faut montrer les deux sens :

- $\Sigma_{mixte} \vdash F \rightarrow x \Rightarrow \Sigma_{R|\tilde{R}} \vdash F \rightarrow x$: Ce sens est trivial du fait que les implications obtenues par les règles et propriétés sont inférées par $\Sigma_{R|\tilde{R}}$.
- $\Sigma_{R|\tilde{R}} \vdash F \rightarrow x \Rightarrow \Sigma_{mixte} \vdash F \rightarrow x$: Ce sens se déduit directement des lemmes 1 et 2.



Le théorème 20 montre que nous pouvons générer une base d'implications mixtes valide et complète d'un contexte réduit R étant données ses bases d'implications positives et négatives.

7.3.4 Exemple complet

Nous reprenons ici l'exemple 27. Nous considérons ainsi la base d'implications génériques mixtes associée au contexte formel R . Pour chacune des implications de cette base, nous associons les implications et les différentes propriétés et règles d'inférence qui nous permettent d'inférer ces implications à partir des bases d'implications génériques positives et négatives associées à R .

$\Sigma_{R \tilde{R}}$	Σ_R	$\Sigma_{\tilde{R}}$	Propriétés/Axiomes
$a\tilde{a} \rightarrow bcde\tilde{b}\tilde{c}\tilde{d}\tilde{e}[0]$			P_2
$b\tilde{b} \rightarrow acde\tilde{a}\tilde{c}\tilde{d}\tilde{e}[0]$			P_2
$c\tilde{c} \rightarrow abde\tilde{a}\tilde{b}\tilde{d}\tilde{e}[0]$			P_2
$d\tilde{d} \rightarrow abce\tilde{a}\tilde{b}\tilde{c}\tilde{e}[0]$			P_2
$e\tilde{e} \rightarrow abcd\tilde{a}\tilde{b}\tilde{c}\tilde{d}[0]$			P_2
$cd \rightarrow abe\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{e}[0]$	$cd \rightarrow abe[0]$		P_1
$ae \rightarrow bcd\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{e}[0]$	$ae \rightarrow bcd[0]$		P_1
$ce \rightarrow abd\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{e}[0]$	$ce \rightarrow abd[0]$		P_1
$de \rightarrow abc\tilde{a}\tilde{b}\tilde{c}\tilde{d}\tilde{e}[0]$	$de \rightarrow abc[0]$		P_1
$\tilde{a}\tilde{b} \rightarrow abcde\tilde{c}\tilde{d}\tilde{e}[0]$		$\tilde{a}\tilde{b} \rightarrow \tilde{c}\tilde{d}\tilde{e}[0]$	P_1
$\tilde{a}\tilde{e} \rightarrow abcde\tilde{b}\tilde{c}\tilde{d}[0]$		$\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	P_1
$\tilde{c}\tilde{b} \rightarrow abde\tilde{a}\tilde{c}\tilde{d}\tilde{e}[0]$		$\tilde{b} \rightarrow \tilde{c}\tilde{d}\tilde{e}$	P_1, P_2
$\tilde{d}\tilde{b} \rightarrow abce\tilde{a}\tilde{c}\tilde{d}\tilde{e}[0]$		$\tilde{b} \rightarrow \tilde{c}\tilde{d}\tilde{e}$	P_1, P_2
$\tilde{e}\tilde{b} \rightarrow abcd\tilde{a}\tilde{c}\tilde{d}\tilde{e}[0]$		$\tilde{b} \rightarrow \tilde{c}\tilde{d}\tilde{e}$	P_1, P_2
$\tilde{c}\tilde{a} \rightarrow abde\tilde{b}\tilde{c}\tilde{d}\tilde{e}[0]$		$\tilde{a} \rightarrow \tilde{c}\tilde{d}$	P_1, P_2
$\tilde{d}\tilde{a} \rightarrow abce\tilde{b}\tilde{c}\tilde{d}\tilde{e}[0]$		$\tilde{a} \rightarrow \tilde{c}\tilde{d}$	P_1, P_2
$b\tilde{c}\tilde{d}\tilde{e} \rightarrow acde\tilde{a}\tilde{b}[0]$		$\tilde{c}\tilde{d}\tilde{e} \rightarrow \tilde{b}$	I_2, P_2
$ab\tilde{c}\tilde{d} \rightarrow cde\tilde{a}\tilde{b}\tilde{e}[0]$			I_6

$\Sigma_{R \tilde{R}}$	Σ_R	$\Sigma_{\tilde{R}}$	Propriétés/Règles d'inférence
$\tilde{e} \rightarrow a$		$\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	I_4
$a \rightarrow \tilde{e}$	$ae \rightarrow bcd[0]$		I_3
$b\tilde{e} \rightarrow a$		$\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	I_4
$ab \rightarrow \tilde{e}$	$ae \rightarrow bcd[0]$		I_3
$\tilde{c}\tilde{e} \rightarrow a$		$\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	I_4
$a\tilde{c} \rightarrow \tilde{e}$	$ae \rightarrow bcd[0]$		I_3
$\tilde{d}\tilde{e} \rightarrow a$		$\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	I_4
$a\tilde{d} \rightarrow \tilde{e}$	$ae \rightarrow bcd[0]$		I_3
$d \rightarrow ab\tilde{c}\tilde{e}$	$d \rightarrow ab$ $cd \rightarrow abe[0]$ $de \rightarrow abc[0]$		P_1 I_3 I_3
$b\tilde{c}\tilde{e} \rightarrow ad$	$d \rightarrow ab$	$\tilde{c}\tilde{d}\tilde{e} \rightarrow \tilde{b}$	I_2 P_1
$ab\tilde{c} \rightarrow d\tilde{e}$			I_6
$c \rightarrow ab\tilde{d}\tilde{e}$	$c \rightarrow ab$ $ce \rightarrow abd[0]$	$\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	P_1 I_3 I_4
$b\tilde{d}\tilde{e} \rightarrow ac$		$\tilde{c}\tilde{d}\tilde{e} \rightarrow \tilde{b}$ $\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{e}[0]$	I_2 I_4
$ab\tilde{d} \rightarrow c\tilde{e}$			I_6
$\tilde{a} \rightarrow be\tilde{c}\tilde{d}$		$\tilde{a} \rightarrow \tilde{c}\tilde{d}$ $\tilde{a}\tilde{b} \rightarrow \tilde{c}\tilde{d}\tilde{e}[0]$ $\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	P_1 I_4 I_4
$e \rightarrow b\tilde{a}\tilde{c}\tilde{d}$	$e \rightarrow b$ $ae \rightarrow bcd[0]$ $ce \rightarrow abd[0]$ $de \rightarrow abc[0]$		P_1 I_3 I_3 I_3
$b\tilde{c}\tilde{d} \rightarrow e\tilde{a}$			I_6
$\tilde{b} \rightarrow a\tilde{c}\tilde{d}\tilde{e}$		$\tilde{b} \rightarrow \tilde{c}\tilde{d}\tilde{e}$ $\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	P_1 I_4
$a\tilde{c}\tilde{d} \rightarrow \tilde{b}\tilde{e}$			I_6
$\tilde{c}\tilde{d}\tilde{e} \rightarrow a\tilde{b}$		$\tilde{c}\tilde{d}\tilde{e} \rightarrow \tilde{b}$ $\tilde{a}\tilde{e} \rightarrow \tilde{b}\tilde{c}\tilde{d}[0]$	P_1 I_4

7.4 Conclusion

Dans ce chapitre, nous avons étudié le problème de génération des implications mixtes d'un contexte formel à partir des bases d'implications génériques positives et négatives qui lui sont associées. Nous avons ainsi exhibé différentes propriétés et règles d'inférence qui nous permettent de générer des implications mixtes à partir de ces données. Malheureusement cet ensemble d'implications mixtes ne suffit pas à définir une base d'implications mixtes complètes pour ce contexte formel. Cette non complétude s'explique en partie par le fait qu'il est impossible de calculer une base d'implications mixtes seulement à partir des bases d'implications positives et négatives, ces bases pouvant correspondre à différents contextes formels. Une suite logique de ce travail est de faire une étude qualitative et quantitative sur l'ensemble des implications que nous pouvons générer par rapport à la base d'implications complète recherchée. Cela permettrait ainsi de voir si l'ensemble des implications que nous générons sont pertinentes. Une autre question reste ici en suspens : Existe-t-il d'autres règles ou propriétés permettant de déduire d'autres règles mixtes à partir des implications positives et négatives ?

D'autre part, nous avons considéré le même problème en ajoutant la contrainte que le contexte considéré est réduit. Cette considération nous a permis de trouver une nouvelle règle d'inférence permettant de déduire de nouvelles implications mixtes. Ce dernier, couplé avec les différentes propriétés et règles d'inférence trouvées ainsi que les axiomes d'Armstrong, nous a permis de montrer qu'il est possible de générer une base d'implications mixtes valide et complète pour un contexte formel réduit étant données ses bases d'implications génériques positives et négatives associées. Nous avons dû pour cela traiter le contexte complet comme un cas particulier.

Ayant ce résultat, on peut se demander s'il n'existe pas des propriétés ou règles d'inférence qui nous permettent de définir une base d'implications négatives d'un contexte réduit, à partir de sa base d'implications génériques positives associée. Si tel était le cas, cela signifierait qu'il est possible de générer une base d'implications mixtes valide et complète pour un contexte formel réduit seulement à partir de sa base d'implications positives. Enfin, est-ce-que toutes les implications positives et négatives sont nécessaires pour pouvoir définir une base d'implications mixtes ? La même question peut ainsi se poser sur les différentes règles d'inférence que nous avons proposées.

Cinquième partie

Conclusion

Chapitre 8

Conclusion et Perspectives

Dans la troisième partie de cette thèse (Différentes opérations sur les systèmes de fermeture), nous avons étudié trois opérations qui sont la borne inférieure de systèmes de fermeture, la borne supérieure de systèmes de fermeture et la différence de systèmes de fermeture. Ces différentes opérations ont été étudiées sur différentes représentations des systèmes de fermeture. Les éléments inf-irréductibles, les bases d'implications et l'opérateur de fermeture associé à un système de fermeture les constituent.

Cette étude a été menée en faisant le lien avec des opérations équivalentes en logique propositionnelle sur les théories de Horn. Nous avons ainsi retranscrit sur les systèmes de fermeture les résultats obtenus sur les théories de Horn et apporté de nouveaux éléments sur ces différentes opérations.

Dans le cas de l'opération de borne inférieure sur les systèmes de fermeture, nous concluons qu'il est facile de la réaliser si nous considérons la représentation par des bases d'implications. En revanche, cette opération s'avère délicate si nous considérons la représentation des systèmes de fermeture par les inf-irréductibles puisque ce problème ne possède pas d'algorithme en temps total polynomial hormis si $P = NP$. De plus, nous proposons un opérateur de fermeture basé sur les opérateurs de fermeture de chaque système de fermeture.

Concernant l'opération de borne supérieure de systèmes de fermeture, les difficultés ne se trouvent pas sur les mêmes représentations que pour l'opération précédente. Ainsi, cette opération est réalisable en temps polynomial si nous considérons la représentation par les éléments inf-irréductibles alors qu'elle ne l'est pas pour les bases d'implications dans le cas général. La considération d'une classe particulière de bases d'implications nous a permis d'aboutir à un résultat intéressant. Nous proposons un algorithme polynomial en la taille des données permettant de résoudre l'opération de borne supérieure de systèmes

de fermeture représentés sous forme de bases d'implications directes.

D'un point de vue pratique, il serait intéressant de tester les performances de cet algorithme en considérant les critères de temps et d'espace, et de les comparer avec celles des algorithmes existants. D'autre part, l'inconvénient des bases d'implications directes réside dans le fait qu'elles peuvent être beaucoup plus grande (taille exponentielle), en nombre d'implications, qu'une base minimum (base canonique de Guigues Duquenne). En revanche, elles possèdent des propriétés algorithmiques que n'a pas une base minimum (exemple : calcul de la fermeture d'un ensemble en un seul passage de la base d'implications). Trouver une base d'implications "hybride" entre les bases d'implications directes et la base d'implications minimum pourrait ainsi permettre de résoudre des problèmes plus efficacement tout en ayant une taille des ensembles d'implications proche de celle de la base d'implications minimum.

Pour ce qui est de l'opération de la différence de systèmes de fermeture, nous avons retranscrit les résultats provenant de la logique propositionnelle. Cette opération n'étant pas propre aux systèmes de fermeture, il nous a paru intéressant de l'aborder puisqu'elle peut aboutir à beaucoup d'applications dans divers domaines. Nous avons ainsi remarqué des résultats assez similaires à l'opération de borne supérieure. L'opération est ainsi réalisable en temps polynomial si nous considérons la représentation par les éléments inf-irréductibles alors qu'elle ne l'est pas pour les bases d'implications. Une perspective de travail est alors de considérer, de la même manière que pour l'opération de borne supérieure, l'opération de différence de systèmes de fermeture représentés par des bases d'implications directes. Permettraient-elles d'aboutir à un algorithme polynomial ?

Dans la quatrième partie de ce manuscrit, nous nous sommes intéressés au problème de génération d'une base d'implications mixtes associée à un contexte formel à partir de ses bases d'implications génériques positives et négatives. Le travail que nous avons réalisé nous a permis d'obtenir des propriétés et des règles d'inférence caractérisant un ensemble d'implications mixtes à partir des implications positives et négatives. De plus, nous sommes arrivés à la conclusion qu'il n'est pas possible de générer une base d'implications mixtes d'un contexte uniquement à partir de ses bases d'implications positives et négatives. Nous avons alors considéré le cas où le contexte pris en compte est réduit. Dans ce cas particulier, une règle d'inférence est exhibée et permet, à l'aide des autres règles et propriétés, de générer une base d'implications mixtes valide et complète à partir des bases d'implications positives et négatives. Une suite logique de ce travail est de faire une étude qualitative et quantitative sur l'ensemble des implications mixtes que nous pouvons générer par rapport à la base d'implications complète recherchée pour le cas d'un

contexte quelconque. En vue de ces résultats, est-ce-que toutes implications positives et négatives sont nécessaires pour pouvoir définir une base d'implications mixtes ? Cette question amène la suivante : Peut-on générer les implications négatives à partir des implications positives à partir des règles d'inférence existantes ou d'autres règles d'inférence ? Enfin, une étude pourrait être faite sur les règles d'association en base de données, qui prendrait en compte la notion de support. Pour ce même domaine, il serait intéressant de caractériser des contraintes permettant d'élaguer de la génération certaines règles qui ne sont pas pertinentes et ainsi réduire la quantité de règles générées qui peut être très importante.

Bibliographie

- [1] R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large database. In *ACM SIGMOD conf. Management of data*, pages 265–290, 1993.
- [2] R. Agrawal and R. Srikant. Fast algorithms for mining association rules. pages 487–499, September 1994.
- [3] L. Nemmiche Alachaher and S. Guillaume. Variables interaction for mining negative and positive quantitative association rules.
- [4] G. Alexe, S. Alexe, Y. Crama, S. Foldes, P. L. Hammer, and B. Simeone. Consensus algorithms for the generation of all maximal bicliques. *Discrete Applied Mathematics*, 145(1) :11–21, december 2004.
- [5] William Ward Armstrong. Dependency structures of data base relationships. In *IFIP Congress*, pages 580–583, 1974.
- [6] W.W. Armstrong. Dependency structures of data base relationships. *Inform. Process. Letter*, 74 :580–583, 1974.
- [7] D. Avis and K. Fukuda. Reverse search for enumeration. *Discrete Applied Mathematics*, 65 :21–46, 1996.
- [8] M. Barbut and B. Monjardet. *Ordre et classification*. Hachette, Paris, 1970.
- [9] C. Beeri and P.A. Bernstein. Computational problems related to the design of normal form relational schemas. *ACM Trans. on database systems*, 1(4) :30–59, 1979.
- [10] A. Berry, A. Sigayret, and J.P. Bordat. Concepts can't afford to stammer. In *Proceedings JIM'03 (Journées de l'Informatique Messine)*, 2003.
- [11] K. Bertet. *Sur quelques aspects algorithmiques et structurels des treillis*. PhD thesis, LIAFA - Université Paris 7, décembre 1998.
- [12] K. Bertet, R. Medina, L. Nourine, and O. Raynaud. Algorithmique combinatoire dans les bases de données massives. In *Acte de l'atelier "Usage des treillis de Galois pour l'intelligence artificielle", conférence AFIA03*, 2003.

- [13] K. Bertet and B. Monjardet. The multiple factes of the canonical direct basis. 2007.
- [14] G. Birkhoff. *Lattice theory*, volume XXV. American Mathematical Colloquium Publications, Providence, third edition, 1967.
- [15] A. Blake. *A Canonical expressions in Boolean algebra*. PhD thesis, University of Chicago, 1937.
- [16] G. Bordalo and B. Monjardet. Reducible classes of finite lattices. *Order*, 13 :379–390, 1996.
- [17] G. Bordalo and B. Monjardet. The lattice of strict completions of a finite poset. *Algebra Universalis*, 47 :183–200, 2002.
- [18] J. P. Bordat. Calcul pratique du treillis de galois d’une correspondance. *Math. Sci. Hum.*, 96 :31–47, 1986.
- [19] J. F. Boulicaut, A. Bykowski, and B. Jeudy. Towards the tractable discovery of association rules with negations. In *FQAS 2000*, pages 425–434, 2000.
- [20] S. Brin, R. Motwani, J. Ullman, and S. Tsur. Dynamic itemset counting and implication rules for market basket data. In *Proc. of ACM SIGMOD 1997*, May 1997.
- [21] Marco Cadoli and Francesco M. Donini. A survey on knowledge compilation. *AI Commun.*, 10(3-4) :137–150, 1997.
- [22] Y. Cai and M.C. Kong. Generating all maximal cliques and related problems for certain perfect graphs. *CONGRESSUS NUMERANTIUM*, 90 :33–55, 1992.
- [23] N. Caspard. *Etude Structurelle et Algorithmique de Classes de Treillis Obtenus par Duplications*. PhD thesis, Université Paris I - Panthéon Sorbonne, 1992.
- [24] N. Caspard, B. Leclerc, and B. Monjardet. *Ensembles ordonnés finis : concepts, résultats et usages*. Springer, 2007.
- [25] N. Caspard and B. Monjardet. The lattices of moore families and closure operators on a finite set : A survey. *Electronic Notes in Discrete Mathematics*, 2 :1–26, April 2000.
- [26] N. Caspard and B. Monjardet. The lattices of closure systems, closure operators, and implicational systems on a finite set : a survey. *Discrete Applied Mathematics*, 127(2) :241–269, April 2003.
- [27] M. Chein. Algorithme de recherche de sous-matrice première d’une matrice. *Bull. Math. R. S. Roumanie*, 13, 1969.
- [28] B.A. Davey and H.A. Priestley. *Introduction to Lattices and Order, Second Edition*. Cambridge University Press, 2002.

- [29] A. Day. A simple solution of the word problem for lattices. *Canad. Math. Bull*, 13 :253–254, 1970.
- [30] R. Dechter and J. Pearl. Structure identification in relational data. *Artificial Intelligence*, (58), 1992.
- [31] Rina Dechter and Judea Pearl. Structure identification in relational data. *Artificial Intelligence*, 58(1-3) :237–270, 1992.
- [32] Alvaro del Val. Approximate knowledge compilation : The first order case. In *AAAI/IAAI, Vol. 1*, pages 498–503, 1996.
- [33] J. Demetrovics, L. Libkin, and I.B. Muchnik. Functional dependencies in relational databases : A lattice point of view. *Discrete Applied Mathematics*, 40(2) :155–185, 1992.
- [34] V. M.F Dias, C. M.H. de Figueiredo, and J. L. Szwarcfiter. Generating bicliques of a graph in lexicographic order. *Theoretical Computer Science*, 37(1–3), June 2005.
- [35] E. Diday and R. Emilion. Maximal and stochastic galois lattices. *Discrete Applied Mathematics*, 127 :271–284, 2003.
- [36] V. Duquenne. The core of finite lattice. *Discrete Mathematics*, 88 :133–147, 1991.
- [37] V. Duquenne. Latticial structures in data analysis. *Theoretical Computer Science*, 217(2) :407–436, 1999.
- [38] V. Duquenne and Guigues. Famille minimale d’implications informatives resultant d’un tableau de données binaires. *Mathématiques et sciences humaines*, 24 :1–7, 1986.
- [39] T. Eiter and G. Gottlob. Identifying the minimal transversals of a hypergraph and related problems. *SIAM Journal on Computing*, 24(6) :1278–1304, 1995.
- [40] T. Eiter and G. Gottlob. Hypergraph transversal computation and related problems in logic and ai. In *European Conference on Logics in Artificial Intelligence (JELIA ’02)*, pages 549–564, 2002.
- [41] T. Eiter, T. Ibaraki, and K. Makino. Disjunctions of horn theories and their cores. Technical report, Rutcor, 1998.
- [42] T. Eiter, T. Ibaraki, and K. Makino. Computing intersections of horn theories for reasoning with models. *Artificial Intelligence*, 110 :57–101, 1999.
- [43] Thomas Eiter, Georg Gottlob, and Kazuhisa Makino. New results on monotone dualization and generating hypergraph transversals. *SIAM J. of Computing*, 32(2) :514–537, 2003.

- [44] Thomas Eiter, Toshihide Ibaraki, and Kazuhisa Makino. On the difference of horn theories. *Lecture Notes in Computer Science*, 1563 :467–477, 1999.
- [45] Thomas Eiter and Kazuhisa Makino. New results for horn cores and envelopes of horn disjunctions. In *Proc. of ECAI2008*, 2008.
- [46] R. Emilion, G. Lambert, and G. Lévy. Algorithms for general galois lattice building. Technical report, CERIA, Université Paris IX Dauphine, 2001.
- [47] M.L. Fredman and L. Kachiyan. On the complexity of dualization of monotone disjunctive normal forms. *Journal of Algorithms*, 21(3) :618–628, 1996.
- [48] H. Fu. *Algorithmique des Treillis de concepts : Application à la fouille de données*. PhD thesis, Université d’Artois, Février 2005.
- [49] H. Fu, H. Fu, P. Njiwoua, and E. Mephu Nguifo. A comparative study of fca-based supervised classification algorithms. In *Proc. of the Second International Conference on Formal Concept Analysis (ICFCA’04)*, LNCS, pages 313–320, 2004.
- [50] H. Fu and E. Mephu Nguifo. A parallel algorithm to generate formal concepts for large data. In *Proc. of the Second International Conference on Formal Concept Analysis (ICFCA’04)*, LNCS, pages 394–401, 2004.
- [51] B. Ganter. Two basic algorithms in concept analysis. Technical report, Technische Hochschule Darmstadt, 1984.
- [52] B. Ganter. Finding closed set under symmetry. Preprint, 1990.
- [53] B. Ganter and K. Reuter. Finding all closed sets : a general approach. *Order*, 8, 1991.
- [54] B. Ganter and R. Wille. *Formal Concept Analysis, Mathematical Foundations*. Springer-Verlag Berlin, 1996.
- [55] V. Garg. Algorithmic combinatorics based on slicing posets. In *Proceedings of the 22th Conference on the Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, december 2002.
- [56] A. Gély. A generic algorithm for generating closed sets of a binary relation. In *Formal Concept Analysis, Proceedings of ICFCA’05*, LNAI, pages 223–234, 2005.
- [57] A. Gély, R. Medina, L. Nourine, and Y. Renaud. Uncovering and reducing hidden combinatorics in guigues-duquenne bases. In *Formal Concept Analysis, Proceedings of ICFCA’05*, LNAI, pages 235–248, 2005.
- [58] M.J. Kearns H.A. Kautz and B. Selman. Reasoning with characteristic models. In *Proc. AAAI-93*, page 34, 1993.

- [59] M. Habib, R. Medina, L. Nourine, and G. Steiner. Efficient algorithms on distributive lattices. *Discrete Applied Mathematics*, 110 :169–187, 2001.
- [60] A. Horn. On sentences which are true of direct unions of algebras. *Journal of Symbolic Logic*, (16) :14–21, 1951.
- [61] D. S. Johnson, M. Yannakakis, and C. H. Papadimitriou. On generating all maximal independent sets. *Information Processing Letters*, 27 :119–123, March 1988.
- [62] M. J. Jou and G. J. Chang. Algorithmic aspects of counting independent sets. *Ars. Combin*, 65 :265–277, 2002.
- [63] T. kashiwabara, S. Masuda, K. Nakajima, and T. Fujisawa. Generation of maximum independent sets of a bipartite graph and maximum cliques of a circular-arc graph. *Journal of Algorithms*, 13 :161–174, 1992.
- [64] R. Khardon. Translating between horn representations and their characteristics models. *Journal of Artificial Intelligence Research*, 3 :349–372, 1995.
- [65] R. Kumar, P. Raghavan, S. Rajapogalan, and A. Tomkins. Extracting large-scale knowledge bases from the web. In *Proc. of the 25th VLDB Conference, Edinburg, Scotland*, 1999.
- [66] R. Kumar, P. Raghavan, S. Rajapogalan, A. Tomkins, and E. Upfal. The web as a graph. In *ACM PODS Conference, Dallas*, 2000.
- [67] S. Kuznetsov. On computing the size of a lattice and related decision problems. *Order*, 18 :313–321, 2001.
- [68] S. Kuznetsov and S. Obiedkov. Comparing performance of algorithms for generating concept lattices. *Journal of Experimental and Theoretical Artificial Intelligence (JETAI)*, 14(2/3) :189–216, 2002.
- [69] S. O. Kuznetsov. On the intractability of computing the duquenne-guigues base. *The Journal of Universal Computer Science*, 10(8) :927–933, 2004.
- [70] C. Lindig. Fast concept analysis. In G. Stumme, editor, *Working with Conceptual Structures - Contributions to ICCS 2000*, pages 235–248, 2000.
- [71] D. Maier. *The theory of relational data bases*. Computer Science Press, Rockville, 1983.
- [72] David Maier. *The Teory of Relational Databases*. Computer Science Press, 1983.
- [73] K. Makino and T. Uno. New algorithms for enumerating all maximal cliques. In *Proc. 9th Scand. Worksh. Algorithm Theory (SWAT 2004)*, LNCS, pages 260–272, 2004.

- [74] H. Mannila and K.-J. Räihä. On the complexity of inferring functional dependencies. *Discrete Applied Mathematics*, 40(2) :237–243, 1992.
- [75] G. Markowsky. Primes, irreducibles and extremal lattices. *Order*, 9 :265–290, 1992.
- [76] R. Medina and L. Nourine. Clone items : a pre-processing information for knowledge discovery, submitted.
- [77] R. Medina, C. Noyer, and O. Raynaud. Efficient algorithm for clones items detection. In *Proc. of the 3rd International Conference on Concept Lattices and Their Applications (CLA'05)*, pages 70–81. ISBN 80-248-0863-3, Palacký University, Olomouc, and VSB - Technical University of Ostrava, 2005.
- [78] R. Missaoui, L. Nourine, and Y. Renaud. Generating positive and negative association rules using formal concept analysis : Problems and solutions. In *Proc. of the Seventh International Conference on Formal Concept Analysis (ICFCA'08)*, pages 169–181, 2008.
- [79] S. Mohseni-Zadeh, P. Brézellec, and J.-L. Risler. Cluster-c, an algorithm for the large-scale clustering of protein sequences based on the extraction of maximal cliques. *Computational Biology and Chemistry*, 28(3) :352–353, July 2004.
- [80] J. B. Nation and A. Pogel. The lattice of completions of an ordered set. *Order*, 14(1) :1–7, 1997.
- [81] L. Nourine. *Une structuration Algorithmique de la Théorie des Treillis – Habilitation à diriger des recherches*. PhD thesis, Université Montpellier II, 2000.
- [82] L. Nourine and O. Raynaud. A fast algorithm for building lattices. *Information Processing Letters*, 71 :199–204, 1999.
- [83] L. Nourine and O. Raynaud. A fast incremental algorithm for building lattices. *Journal of Experimental and Theoretical Artificial Intelligence*, 14 :217–227, 2002.
- [84] C.H. Papadimitriou. *Computational Complexity*. Addison Wesley, 1994.
- [85] N. Pasquier, Y. Bastide, T. Taouil, and L. Lakhal. Efficient Mining of Association Rules Using Closed Itemset Lattices. *Information Systems*, 24(1) :25–46, 1999.
- [86] J. Pei, J. Han, and R. Mao. Closet : An efficient algorithm for mining frequent closed itemsets. In *Proc. of ACM-SIGMOD, International Workshop on Data Mining and Knowledge Discovery (DMKD'00)*, Dallas, May 2000.
- [87] G. Pruesse and F. Ruskey. Gray codes from antimatroids. *Order*, 10 :239–252, 1993.
- [88] W. Quine. A way to simplify truth functions. *American Mathematical Monthly*, 62 :627–631, 1955.

- [89] Yoan Renaud. Join on closure systems using direct implicational basis representation. 2008.
- [90] A. Savasere, E. Omiecinski Shamkant, and B.Navathe. Mining for strong negative associations in large database of customer transactions.
- [91] B. Selman and H. Kautz. Knowledge compilation using horn approximations. In *Proc. AAAI-91*, page 904, 1991.
- [92] R. C. Shock. Computing the minimum cover of functional dependencies. *Information Processing Letters*, 3(22) :157–159, 1986.
- [93] Einoshin Suzuki. Data mining methods for discovering interesting exceptions from an unsupervised table. *Journal of Universal Computer Science*, 12(6) :627–653, 2006.
- [94] S. Tsukiyama, M. Ide, M. Aiyoshi, and I. Shirawaka. A new algorithm for generating all the independent sets. *SIAM J. Computing*, 6 :505–517, december 1977.
- [95] T. Uno. A new approach for speeding up enumeration algorithms and its application for matroid bases. In *Proceedings of Computing and Combinatorics : 5th Annual International Conference, COCOON'99, Tokyo, Japan*, LNCS, July 1999.
- [96] T. Uno, T. Asai, Y. Uchida, and H. Arimura. Lcm : An efficient algorithm for enumerating frequent closed item sets. In *ICDM'03 - Proc. of Workshop on Frequent Itemset Mining Implementations (FIMI)*, 2003.
- [97] T. Uno, M. Kiyomi, and H. Arimura. Lcm ver. 2 : Efficient mining algorithms for frequent/closed/maximal itemsets. In *ICDM'04 - Proc. of Workshop on Frequent Itemset Mining Implementations (FIMI)*, 2004.
- [98] P. Valtchev, R. Missaoui, and P. Lebrun. Data mining using nested lattices. Technical Report 2000-08, Dept of CS, University of Quebec in Montreal, July 2000.
- [99] J. Wang, J. Han, and J. Pei. Closet+ : Searching for the best strategies for mining frequent closed itemsets. In *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Washington, D.C., August 2003.
- [100] M. Wild. Implicational bases for finite closure systems. Preprint, 1989.
- [101] M. Wild. Optimal implicational bases for finite modular lattices. *Quaestiones Mathematicae*, 23 :153–161, 2000.
- [102] Xindong Wu, Chengqi Zhang, and Shichao Zhang. Efficient mining of both positive and negative association rules. *ACM Trans. Inf. Syst.*, 22(3) :381–405, 2004.
- [103] M. J. Zaki and C.-J. Hsiao. Charm an efficient algorithm for closed itemset mining. In *2nd SIAM International Conference on Data Mining, Arlington*, April 2002.

- [104] M. J. Zaki and C.-J. Hsiao. Efficient algorithms for mining closed itemsets and their lattice structure. *IEEE Transactions on knowledge and data engineering*, 17(4) :462–478, April 2005.
- [105] M. J. Zaki and M. Ogihara. Theoretical foundations of associations rules. In *3rd SIGMOD'98 Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD)*, Seattle, WA, pages 7 :1–7 :8, June 1998.

Nous présentons dans cette thèse les définitions et notations liées aux systèmes de fermeture et montrons leur relation avec les théories de Horn. Nous nous intéressons ensuite à trois opérations sur les systèmes de fermeture: la borne supérieure, la borne inférieure et la différence. Nous proposons une caractérisation de ces différentes opérations selon la représentation des systèmes de fermeture que nous considérons.

On s'intéresse ensuite au problème de génération d'une base d'implications mixtes d'un contexte formel. Nous étudions ce problème lorsque la donnée prise en considération est constituée des bases d'implications génériques positives et négatives de ce contexte. Trois résultats majeurs sont présentés: l'apport de propriétés et de règles d'inférence pour déduire des implications mixtes, l'impossibilité de générer une base d'implications mixtes juste et complète à partir de ces données dans le cas général, et la faisabilité dans le cas où le contexte est considéré réduit.

This thesis deals with some algorithmic aspects on closure systems and their different representations. We present different definitions and notations on closure systems and the link between closure systems and Horn theories, we are interested in three operations on closure systems: join, meet and difference. This thesis propose a characterization on this operation with respect to representation of closure systems we consider. We translate results of equivalent problems studied on Horn theory and give some new contributions.

Then, we deal with generation of a mixte implicational basis of a formal context. We are particularly interested in this problem where the input is positive and negative implicational basis of this context. Three results are presented: properties and rules are given to find mixte implications, the impossibility of generating a complete mixte implicational basis having just this data in general case and feasibility in case where context is considered reduced.